Variationally Computed Line Lists for SO_2 and SO_3

Daniel Underwood

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Physics & Astronomy

University College London

April 19, 2016

I, Daniel Underwood, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Associated Publications

1. **An ab initio variationally computed room-temperature line list for $^{32}S^{16}O_3$**
   Daniel S. Underwood, Sergei N. Yurchenko, Jonathan Tennyson
   Phys. Chem. Chem. Phys., **15**, 10118-10125 (2013); DOI: 10.1039/C3CP50303H

2. **Rotational spectrum of SO_3 and theoretical evidence for the formation of sixfold rotational energy-level clusters in its vibrational ground state**
   Daniel S. Underwood, Sergei N. Yurchenko, Jonathan Tennyson, Per Jensen
   J. Chem. Phys., **140**, 244316 (2014); DOI: 10.1063/1.4882865

Abstract

The work presented in this thesis concerns the production of high-temperature spectroscopic line lists for the SO_2 and SO_3 molecules, for the purposes of astrophysical, terrestrial and industrial applications. Both line lists are computed using *ab initio* computational methods to calculate rovibrational energy levels and dipole moment transition intensities.

The hot SO_2 line list is computed using the DVR3D program suite, optimised to work efficiently with the molecule, and by making use of high performance computing. The calculations are based on an empirically refined *ab initio* potential energy surface (PES), and a purely *ab initio* dipole moment surface (DMS). Results are compared to previous *ab initio* studies and available experimental data. The final line list can be used in spectroscopic models for temperatures up to and including 1500 K.

A preliminary, room-temperature line list for SO_3 is calculated using the TROVE program, in conjunction with a purely *ab initio* PES and DMS. The results are compared to available experimentally derived energy level data. These are then used to empirically refine the *ab initio* PES, which is subsequently employed in the calculation of the complete, hot line list, suitable for modelling SO_3 spectra up to 773.15 K. Preliminary comparisons are made with experimental high-temperature measurements, and the quality of the *ab initio* DMS is discussed.

In addition, the rotational behaviour of the SO_3 molecule is investigated from a theoretical perspective using the synthetic SO_3 line list, where the 'forbidden' rotational spectrum is analysed. The formation of so-called 6-fold rotational energy clusters at high rotational excitation is also predicted, the dynamics of which are analysed both quantum mechanically and semi-classically.

The potential applications and limitations of both line lists is outlined, and implications for further work are also discussed.

Acknowledgements

I would like to thank my supervisors Prof. Jonathan Tennyson and Sergey Yurchenko for their excellent support and advice throughout the duration of this work. Their unequalled expertise provided the foundation to my efforts and acquired knowledge, and their continuous encouragement and guidance made achieving the results of this thesis possible.

I would also like to thank my parents for their never-ending love and support, and for having the patience to put up with me in my exhausting attempt to complete this thesis; to Kerry O'Callaghan for providing a huge amount of emotional support and respite; to my colleagues at UCL, within the ExoMol group, and beyond, including George Kelly, David Johnson, David Buckley, Richard Rollins and Paul Moseley; to my various housemates and friends who I have had the pleasure of knowing throughout my time in London, including Tom Stanton, Sam Maclean, Joshuha Thomas-Wilsker, Will Dunn, Megan Nee, Emily Morrison, and the rest of the Streatham Social Club.

Thanks must also be given to Alexander Fateev and his team at the Danish Technical University, for stimulating the work presented in this thesis, and for providing a view into the experimentalist's world; to Xinchuan Huang and Timothy Lee at NASA Ames for providing the necessary material for computing the SO_2 line list, and their continuing support throughout the investigation; to Jeff Barber for supplying experimental data for SO_3; to Ahmed Al-Refaie, who provided the necessary algorithm which made the completion of the SO_3 line list possible.

This work was supported by grant 10442 from energinet.dk under a subcontract from the Danish Technical University and the ERC under Advanced Investigator Project 267219.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The understanding of sulphur chemistry is of great interest for a variety of astrophysical, terrestrial and industrial environments. The cosmic abundance of sulphur and its significant role in the current comprehension of biological processes mean that its importance cannot be overstated. Sulphur exists in many forms, one of which is the molecular compound form of the sulphur oxides. There are many stable forms of such compounds, although arguably the most important of these are sulphur dioxide (SO_2) and sulphur trioxide (SO_3). In their gaseous states their presence is well-known in a variety of settings, and their coexistence is often linked via the same chemical processes. For these reasons the detection and monitoring of these species is of particular importance across a large number of scientific disciplines.

The analysis of molecular spectra in conjunction with quantum mechanics is essential in the understanding of the physical properties of molecules, and the field of molecular spectroscopy is varied and well-established. The characteristics of observed spectra provide insight into the structure and motional behaviour of molecules, as well as the physical properties of their environment; in particular, infrared and microwave spectroscopy concern the study of the vibrational and rotational motions of molecules. The discrete nature of energy states associated with these motions for a particular molecule, and the transitions between them due to their interaction with electromagnetic radiation give rise to the well-known features within these spectral regions. A firm understanding and interpretation of the unique spectral characteristics of a molecule can therefore be exploited to determine its presence, abundance and behaviour in a large number of remote physical settings, as well as providing a means to probe the conditions of such environments.

SO_2 and SO_3 are certainly no exception. The work described in this thesis concerns the production of an extensive set of spectroscopic data for both SO_2 and SO_3 from a theoretical perspective, in order to supplement observational analyses pertaining to these molecules. In addition to this, the simulation of the spectral characteristics of these molecules provides an opportunity to rigorously assess already established theoretical models, and augment them appropriately to suit the requirements of the desired results.

In this introductory chapter the importance of the spectroscopic description for both SO_2 and SO_3 is considered within a variety of different contexts, including an outline of the already existing experimental and theoretical work on the subject. The subsequent need to extend upon these studies is also addressed.

1.1 Background & Motivation

The presence and detection of both SO_2 and SO_3 in their gaseous states may be explained by either physical processes that are naturally occurring or by those that are influenced by human activity. There are a number of different isotopologues of both SO_2 and SO_3 which undoubtedly contribute in varying, minor degrees to the sulphur oxide chemistry described in the literature. The most abundant of these are $^{32}S^{16}O_2$ and $^{32}S^{16}O_3$. These main isotopologues are the focus of this thesis, and reference to them is implicit throughout. This section highlights some important settings where both species are known or thought to play a significant role.

1.1.1 SO_2 and SO_3 in Astrophysical & Terrestrial Environments

The study of atmospheric physics and chemistry provides a key insight into the properties of the Earth's atmosphere, and both solar and extra-solar system bodies. The sophisticated modelling and understanding of these environments relies heavily on the knowledge of various spectroscopic processes involving atomic and molecular species that are present. Spectral measurements may be performed in order to gain insight into such processes, via both ground and satellite based observations, as well as in laboratory-based replications.

Both SO_2 and SO_3 are well-known to exist naturally in Earth's atmosphere. The main source of their natural occurrence is via volcanic emissions and hot springs [1–3];

observation of these products provide a useful tool in the understanding of such geological processes. The spectroscopic study of these sulphur oxides can also provide insight into the history of the Earth's atmosphere [4]. However the most important aspect of the existence of these two species is arguably their contribution to the formation of acid rain [5–8]. The oxidisation of SO_2 to SO_3 in the atmosphere, followed by subsequent rapid reaction with water vapour results in the production of sulphuric acid (H_2SO_4), which leads to many adverse environmental effects. The study of these chemical processes is therefore of great importance. Though a product of natural processes, the effect is greatly enhanced by human activity, particularly via byproducts of industrial applications; these are addressed in Subsection 1.1.2.

The remote sensing techniques of the Earth's atmosphere can be extended towards distant astrophysical bodies. In the absence of *in situ* measurements, and not to mention their associated impracticalities, spectroscopy provides the only means of analysing the characteristic behaviours of distant planetary atmospheres, as well as other environments of interest. Sulphur oxide chemistry has been observed in a variety of astrophysical settings. Within the solar system, both SO_2 and SO_3 are thought to be major constituents in the atmosphere of Venus [9–15] and Jupiter's moon, Io [16–19]; all of these are known to have volcanic activity. It has been observed in the atmosphere of Mars, although to a much lesser extent [3]. These molecules are often found alongside sulphur monoxide (SO) [3, 9, 13], which is also an important molecule with regards to sulphur chemistry. The presence of these sulphur oxides has also been attributed to environments beyond the solar system, including the atmospheres of giant planets, brown dwarfs, and dwarf stars [20], circumstellar envelopes of young and evolved stars [21–25], and in molecular clouds and nebulae within the interstellar medium [26–31]. Extragalactic detection of SO_2 has even been achieved [32, 33], emphasising the universal abundance of this particular compound, along with SO [34].

One of the most exciting developments in recent years lies in the discovery of extrasolar planets, or "exoplanets". The observation of the tremendous variety of such bodies has challenged the current understanding of solar system and planetary formation. Exoplanet detection methods have grown in sophistication since the inception of the field, however efforts to characterise their atmospheres are relatively new. The well-documented distribution of sulphur oxides in various terrestrial and astrophysical

environments mean that a thorough understanding of their fundamental spectroscopic behaviour is essential in the future analysis of the spectra of these exoplanetary atmospheres, and of other bodies of interest observed through past, present and future space telescope missions [35, 36].

1.1.2 SO_2 and SO_3 in Industrial Environments

The measurement and calibration of SO_2 and SO_3 spectra is of particular relevance to industrial applications. Their presence is observed in the products of various combustion processes involving sulphur-containing fuel [8, 37, 38], and analyses of these is often useful, for example in monitoring the performance of jet engines [39]. SO_2 and SO_3 are also observed in selective catalytic reduction (SCR) units, where the presence of both is undesirable within flue gas chambers in large quantities, as well as other industrial exhausts [40]. SCR systems are used to reduce NO_x emissions with the aid of a catalyst, producing an array of chemical reactions within gas chambers; the production of SO_2, and SO_3 within secondary reactions, due to a catalytic process, can have damaging effects due to their acidic properties; they are corrosive, making their production hard to monitor [41] and, as previously stated, form a major constituent in the production of acid rain [5]. The control of their output is therefore of great importance.

A team led by Alexander Fateev at the Technical University of Denmark (DTU) have been working to develop *in situ* infrared measurement techniques for a variety of gaseous species at high temperatures, making the need for accurate spectral assignments a priority [42, 43]. Both SO_2 and SO_3 are considered species of interest, due to the requirement of their close monitoring in these industrial settings, and therefore assignments of their spectra at high temperature are required in order for *in situ* monitoring to be suitable for industrial applications on SCR units, with the ultimate aim of optimising the reduction of NO_x emissions and ammonia consumption, and minimising material corrosion to prolong the lifetimes of these units.

The development of this technique has involved the measurement of species concentrations for SO_2 and SO_3 within a hot gas flow cell in laboratory conditions [42], which is essential in the characterisation of the IR spectra of these molecules within flue gases. The temperature ranges of interest lie within 500 - 800 K, however this technique of spectral assignment has its limitations due to the high reactance of these

species. Laboratory production of SO_3 can be achieved by the reaction of SO_2 with either O_2 or O_3, though despite efforts in the design of gas cells optimised for use with reactive gases, corrosion of gas cell components have still been observed, suggesting losses of SO_3/H_2SO_4 between the generation unit and the gas flow cell where IR measurements are performed.

1.1.3 The Requirement of Infrared Data

The importance of SO_2 and SO_3 within the previously described physical environments has meant that they have been the subject of a large number spectroscopic studies (see Section 1.2) which include analyses in both the infrared (IR) and ultraviolet (UV) regions of the electromagnetic spectrum. There are a number of reasons for further extending upon these studies, particularly in the IR, which serve as the justification for the current work.

The UV (electronic) spectrum of SO_2 is easily observed and widely used in the detection of SO_2 in astrophysical and industrial analyses [42, 44–48]. However, despite the numerous studies of the SO_2 UV spectrum in the laboratory, the properties of this molecule in these spectral regions are still poorly understood, and it is difficult to accurately predict its absorption behaviour as a function of temperature [49–52]. Experimental studies of the UV spectrum have of course been complemented by theoretical works [53–57], though they are limited in that they may only draw comparison with room-temperature data. The case is similar for SO_3, which has, in general, been studied less so than SO_2 experimentally [58–60], and not at all theoretically in the UV. Supplementary studies of these molecules in the IR can help paint a clearer picture as to their presence and abundance, where UV observations may be inadequate.

Huang *et al.* [35] described the need for accurate IR SO_2 data by astronomers to effectively analyse astronomical atmospheres and identify other molecular species of interest. A large number of observing instruments provide an insight into the IR spectral range where many important molecules are spectroscopically active, such as the Herschel Space Observatory [61], the Spitzer Space Telescope [62], and the Stratospheric Observatory for Infrared Astronomy [63]. From an alternative point of view, the presence of SO_2 within recorded spectra may be considered a hindrance to the understanding of the astrophysical environments under observation. The problem lies in

the fact that, in efforts to identify important molecular species within various astronomical settings, the presence of extraneous features attributed to other superfluous molecules within these spectral regions makes it particularly difficult to identify those that are of interest. A report from the workshop on laboratory spectroscopy in support of Herschel, SOFIA, and ALMA [64] describes the need of removing spectral 'weeds' in order to allow for the inference of more important species. SO_2 has been described as such a 'weed', and therefore a firm understanding of its IR spectrum at high resolution is required in order to identify and subtract out its influence from recorded data in these regions, and high-temperature data is necessary to account for the increasing effect of interference in warmer regions of interest.

In a similar vein, and more closely related to the aims of the current work, the identification of SO_3 in the IR is hindered by the presence of interfering SO_2 in environments where both species are found simultaneously; a number of their spectral features overlap, particularly the ν_3 bands of both molecules in the 1300 - 1400 cm^{-1} region (see Section 1.2). From this point of view SO_2 can also be seen as a spectral 'weed' with respect to the detection of SO_3. A firm understanding of the spectroscopic behaviour of both of these molecules within the same spectral window is therefore required to be able to correctly identify each species independently.

These reasons, along with the practical problems described in subsection 1.1.2 that make the spectral assignment of each species a difficult task form the motivation for the current work: to obtain an extensive resource of IR spectroscopic data for both SO_2 and SO_3 via *ab initio* computational methods, in order to supplement the characterisation of hot spectra within industrial environments. Such work is performed by the ExoMol group, which is discussed in Section 1.4. First, a brief review is given of the pre-existing experimental and theoretical work carried out in order to achieve such results, and a discussion of molecular spectroscopic databases.

1.2 Previous Spectroscopic Studies on SO_2 and SO_3

Both SO_2 and SO_3 have already been the subject of a wide variety of spectroscopic studies. This section highlights some important relevant works carried out with regards to the spectroscopy of both these molecules, in both experimental and theoretical contexts.

1.2.1 Experimental Studies

1.2.1.1 SO_2

The importance of SO_2 in the various physical and chemical settings previously described means that it has been the focus of a great number of experimental spectroscopic studies, and subject to a number of reviews. Early studies focusing on the structure of the molecule using microwave observations measure the equilibrium S-O bond length at 1.43076 Å and O-S-O bond angle at 119° [65, 66]. Its bent molecular shape gives it the characteristics of $\mathscr{C}_{2v}$ symmetry (see Chapter 2), and a permanent dipole moment of 1.62673 D [67]. The fundamental vibrational frequencies ν_1, ν_2 and ν_3 have been measured as 1151.7, 517.9 and 1362.1 cm^{-1}, respectively [68], and consistently reaffirmed [66]. Spectra have also been analysed for the purposes of measuring absolute intensities [69–71] and determining dipole moments [67, 72, 73]. Spectra have even been obtained from supersonic jet engines [37, 38] for the purposes of vibronic analysis, and phosphorescence studies have also been performed [74]. The range of work is vast; with the advancing improvement of spectroscopic observational tools the infrared spectrum of SO_2 is constantly being reassessed with ever increasing resolution.

For the purpose of brevity it is perhaps best to focus upon infrared studies attributed to aggregated data within the HITRAN database for high-resolution transmission, molecular absorption data [75] (see Section 1.3) for this molecule, for which assigned transition wavenumbers and intensities are provided. There are a total of 14 bands available in the HITRAN database for SO_2, consisting of 72 459 lines for the main $^{32}S^{16}O_2$ isotopologue (there are a further 22 662 for $^{34}S^{16}O_2$).

Purely rotational transitions are provided by the CDMS database [76] (see Section 1.3) for both $\nu_0 \leftarrow \nu_0$ and $\nu_2 \leftarrow \nu_2$. The majority of remaining band data is provided by Lafferty *et al.* for ν_2, $2\nu_2 \leftarrow \nu_2$, $3\nu_2 \leftarrow 2\nu_2$, ν_1, $\nu_1+\nu_2 \leftarrow \nu_2$, ν_3, $\nu_1+\nu_3$, $\nu_1+\nu_2+\nu_3 \leftarrow \nu_2$, $2\nu_3$, $3\nu_3$ and $\nu_1+3\nu_2$ [68, 70, 71, 77, 78], with additional data for ν_3 and $\nu_2+\nu_3 \leftarrow \nu_2$ by Henningsen *et al.* [79]. Additional absolute intensity data is provided in conjunction with corresponding bands from Refs. [69, 80]. The majority of these transitions possess a HITRAN error code for line position and intensity of $ierr \geq 4$, meaning their uncertainties are less than 1×10^{-3} cm^{-1} and 20%, respectively. The exception to these is in the $3\nu_3$, $\nu_1+\nu_3$, and $\nu_1+\nu_2+\nu_3 \leftarrow \nu_2$ bands, for which

error codes are unreported for both line position and absolute intensities. The set of HITRAN data for SO_2 provides the main source of experimental comparison to the results produced in the current work, which is detailed in Chapter 3.

The most recent high resolution analysis of SO_2 has been carried out by Ulenikov *et al.* for a number of bands [81–86]. These publications cover 15 vibrational bands for the purposes of transition and parameter assignment via effective Hamiltonian models. The following bands were analysed: ν_1, ν_3, $3\nu_1$, $2\nu_2$, $3\nu_2$, $2\nu_3$, $\nu_1+\nu_2$, $\nu_1+3\nu_3$, $2\nu_1+\nu_2$, $3\nu_1+\nu_3$, $\nu_2+\nu_3$, $\nu_1+2\nu_2 \leftarrow \nu_2$, $\nu_2+3\nu_3 \leftarrow \nu_2$, $\nu_2+2\nu_3 \leftarrow \nu_2$, $2\nu_2+\nu_3 \leftarrow \nu_2$ and $2\nu_1+\nu_2+\nu_3 \leftarrow \nu_2$. A number of these bands are not covered in the HITRAN database, however absolute intensities were not reported for any of these measurements.

The ultraviolet (electronic) spectrum of SO_2 has also been the subject of spectroscopic interest, and exhibits an easily observed, but complex structure. There are three UV absorption bands within the 170 - 400 nm window [87] which are all relevant to atmospheric studies. The strongest of these is attributed to transitions from the ground X^1A_1 electronic state to the excited $C^1B_2 - X^1A_1$ state, and extends from 170 to 230 nm [49, 87]. A number of cross section measurements have been made for this band [87–90] at various resolutions and temperatures. Between 240 and 340 nm, UV features are mainly associated with the $B^1B_1 - X^1A_1$ band, but include vibronic coupling contributions from the $^3B_1$, $^3B_2$ and $^1A_2$ states [91]. The perturbation of the $B^1B_1 - X^1A_1$ is such that it is difficult to make rovibrational assignments, due to the inability to resolve rotational structure from a dense continuum of weak absorption features [52]; nevertheless, cross sections for this band have been reported in the literature [49, 52, 90, 92], and some vibrational structure analyses have also been performed [50, 93, 94]. Finally, the region 345 - 400 nm contains a weak absorption feature which has been assigned to the $a^3B_1 - X^1A_1$ electronic transition, which is spin-forbidden [87, 95]. The assignment of this band has been attempted with varying results [50, 51], and with some confirmations [38]. The temperature dependence of this band was also analysed [52], which also showed that previous assignments were unsatisfactory.

1.2.1.2 SO_3

The experimental IR spectroscopic study of SO_3 is not as extensive as that of SO_2, which may be attributed to its extreme chemical reactivity which make measurements difficult. Historically vibrational Raman [96], infrared [97] and electron diffraction studies [98] have all contributed to the understanding of the structure of this molecule, as well as its fundamental vibrational frequencies. Its equilibrium position was found to be that of a symmetric planar molecule with equilibrium S-O bond lengths of 1.41732 Å and inter-bond angles of 120° [99], described by $\mathscr{D}_{3h}$(M) symmetry (see Chapter 2), which makes it a suitable example for testing effective Hamiltonian models for these types of symmetric tops. The ν_1, ν_2, ν_3 and ν_4 fundamental frequencies are attributed to the totally symmetric stretch at 1064.9 cm^{-1} [100], the symmetric bend at 497.5 cm^{-1}, and the asymmetric stretching and bending modes at 1391.5 and 530.1 cm^{-1}, respectively [101].

The infrared and coherent anti-stokes vibration-rotation spectra of a number of isotopologues of SO_3 have been extensively investigated in a series of papers by Maki *et al.* [99–105], reassessing and confirming fundamental constants and frequencies. Several bands were analysed based on an empirical fitting to effective Hamiltonian models, yielding rovibrational constants and energy levels assigned by appropriate vibrational and rotational quantum numbers. In total,18 bands were analysed: ν_2, ν_3, ν_4, $2\nu_3^{(l_3=2)}$, $3\nu_3^{(l_3=1)}$, $3\nu_4^{(l_4=1)}$, $\nu_1+\nu_2$, $\nu_1+\nu_4^{(l_4=1)}$, $\nu_2+2\nu_4^{(l_4=0)}$, $\nu_2+2\nu_4^{(l_4=2)}$, $2\nu_2+\nu_4$, $\nu_3+\nu_4$, $\nu_1 \leftarrow \nu_4$, $\nu_2+\nu_4 \leftarrow \nu_2$, $2\nu_2 \leftarrow \nu_2$, $\nu_2+\nu_4 \leftarrow \nu_4$, $2\nu_4^{(l_4=0)} \leftarrow \nu_4$ and $2\nu_4^{(l_4=2)} \leftarrow \nu_4$.

The work also provided some information on transition intensities (for room-temperature), however these were secondary to the main focus of the study, provided only as supplementary material in the form of relative intensity values for a limited number of bands, and at the time of writing are the only known intensities published to the literature. The associated supplementary data of these studies for both the rovibrational energies and relative intensities have been used as a basis for comparison in the current work on SO_3, and their use is described in detail in Chapters 4, 5, and 6.

The "forbidden" rotational spectrum, for which centrifugal distortions can induce transitions, was investigated for the first time by Meyer *et al.* using microwave Fourier-transform spectroscopy [106]. Assignments for 25 transitions were made, as well as

the determination of a number of rotational constants, including the only direct measurement of the C_0 rotational constant.

There have been a few studies on the ultraviolet spectrum of SO_3 by Fajans and Goodeve [58], and LeRoy *et al.* [59], both between 220 and 270 nm where overlap with SO_2 is small. A further study by Burkholder and McKeen [60] reported cross sections for the 195 to 330 nm range for the purposes of photolysis rate calculations of SO_3. All measurements were taken at room-temperature, and neither reported assignments for any of the bands, which exist as weak, diffuse vibrational band structures superimposed on a continuous background. As such, the rovibronic behaviour of SO_3 is much less understood than for SO_2.

1.2.2 Theoretical Studies

1.2.2.1 SO_2

Theoretical IR studies on the spectroscopy of SO_2 have been numerous. The development of potential force fields for the purposes of rovibrational energy calculation has been the subject of many efforts using various *ab initio* methods [107–114], all with varying degrees of accuracy in reproducing experimental observations. For example, Pak and Woods [113] reported equilibrium constants as 1.4417 Å and 118.9° for the S-O and O-S-O bond lengths and angles, respectively, as well as a permanent dipole moment of 1.6909 Debye. This was based upon a CCSD(T) (coupled-cluster with single and double excitations) approach [115] in conjunction with a Gaussian-type basis set (see Chapter 2 for a discussion on basis sets and the coupled-cluster method, along with other electronic structure methods). Martin reported improved values of 1.43043 Å and 119.26° based on the same coupled cluster approach, and a variety of basis sets. The SO_2 fundamentals were reproduced with errors of +3.9, -0.4, and +0.4 cm^{-1}. These works have set the precedent for *ab initio* methods for computations regarding SO_2, however they show a good example of the need to spectroscopically refine potential parameters based on experimental data.

An example of SO_2 IR intensity simulation for high temperatures has been attemptted by Voitsekhovskaya *et al.* for the purposes of analysing jet engine exhausts [39, 116]. These are based on a model Hamiltonians in conjunction with potential parameters for reproducing rotational energy levels, as well as rovibrational levels for

the three fundamental bands. The method for intensity simulation is based upon experimentally derived dipole derivatives, and reproduces integrated intensities at room-temperature reasonably well for these bands. Such methods are not particularly well-suited for high-resolution studies, however, and more sophisticated *ab initio* methods for dipole calculations are required.

A limited IR line list was produced for SO_2 by Fortenberry *et al.* [117] on the basis of quartic force field models to produce transition wavenumbers, and *ab initio* dipole moment surfaces to calculate transition intensities, based on MP2 and coupled clustered methods. The calculation of transition dipoles via this method is necessary for extensive coverage of intensities. This work was drastically expanded upon in a more advanced theoretical spectroscopic work of SO_2 by Huang *et al.* [35] in the calculation of room-temperature, infra-red line lists for a number of SO_2 isotopologues [118]. This study, which was published during the timeframe of the work of this thesis, involved the construction of experimentally refined *ab initio* potential energy and dipole moment surfaces, computed using coupled cluster methods, which then serve as a basis in the nuclear motion calculations and intensity simulations, respectively. A theoretically derived S-O bond length of 1.431086 Å was obtained, as well as 1.629402 D for the permanent dipole moment, both agreeing well with experimentally retrieved values [67]. Comparisons of rovibrational transitions were made with HITRAN data [75] and to the data of Ulenikov *et al.* [81–86] mentioned previously, with root-mean-squares errors ranging from 0.01 - 0.03 cm^{-1}. The reported line lists are suitable for the modelling of spectra in the $0 < \nu \leq 8000$ cm^{-1} region, involving transitions between energy levels up to 13 168 cm^{-1}, for all $J \leq 80$. This work by Huang *et al.* forms part of a collaborative effort in conjunction with the work on SO_2 which is detailed in this thesis, and deals with extending this work via the computation of a hot line list by use of the PES and DMS constructed in Refs. [35]. This is the subject of Chapter 3.

Work has also been done on the UV spectrum of SO_2 from a theoretical perspective, involving the calculation of potential surfaces for the ground X^1A_1 electronic state, and some low-lying excited states; the majority of these deal with the C^1B_2 excited state. For example, Nachtigall *et al.* investigated the stationary points of the C^1B_2 state by means of high-level *ab initio* methods [119] and later used the PES to calculate vibrational levels up to dissociation [120]. Further studies by this group included the

calculation of absorption and resonance emission cross sections involving the X^1A_1 and C^1B_2 states, in conjunction with transition dipole functions [53], giving semi-quantitive agreement with experiment. Similar work has been performed by Ran *et al.* [54] and by Tokue and Nanbu [55]. Lévêque *et al.* [56] provided one of the few examples of *ab initio* work on the coupled $^1B_1$ and $^1A_2$ states, with nearly quantitive agreement with low-resolution experiment. A similar study was conducted by Xie *et al.* [57], additionally providing *ab initio* potential surface calculations of the a^3B_1 and b^3A_2 for the first time.

1.2.2.2 SO_3

There has been limited theoretical work on SO_3. Dorney *et al.* [121] reported the first anharmonic force constants based on a valence force model, with the aim of reproducing a number of well-determined spectroscopic parameters available at the time of publication. The work yielded a value of 1.4184 $\pm$ 0.001 Å for the equilibrium bond length. Flament *et al.* [122] published *ab initio* harmonic force constants for SO_3, calculated via a 'scaled quantum mechanical' method, as well as through the multi configurational self consistent field (MCSCF) method, and computed vibrational frequencies agreeing to within 4% of the established experimental values.

The first accurate, fully *ab initio* anharmonic quartic force field for $^{32}S^{16}O_3$ was published by Martin [123]. This work involved computation using coupled cluster methods (CCSD(T)) [115] in conjunction with a number of different basis sets, and reported theoretical estimates for the band origins of the low-lying vibrational states. The results reproduced the equilibrium bond length to within +0.0003 Å, and the fundamentals to within 1.15 cm^{-1}, on average.

As in the case of experimental studies for SO_3, there exist no absolute transition intensity values derived from theoretical approaches. The current work provides the first instance of such an effort, and is the subject of Chapters 4 and 6. There have also been no theoretical studies into the UV spectrum of SO_3.

1.3 Spectroscopic Databases

The large number of studies on SO_2 and SO_3, and indeed every molecular species of interest, provides a huge, multi-sourced set of spectroscopic data from which to choose from. For the purposes of atmospheric modelling, quantum mechanical analysis, and

other situations which require the use of spectroscopic parameters, it is often useful to have a standardised set of reliable data to work with.

The aggregation of molecular spectroscopic data is a familiar concept, as can be seen in the variety of pre-existing spectroscopic databases, such as HITRAN [75], GEISA [124, 125], CDMS [76], NIST and PNNL [126], and JPL [127]. These databases provide experimentally accurate data for a variety of molecular species, obtained from a number of sources. For example HITRAN (high-resolution transmission) [75] was originally set up to provide parameters for the simulation of atmospheric transmission of infrared light based on a limited number of important atmospheric absorbers on Earth, but has since expanded to include a larger spectral range for a greater number of molecules important in a wider range of environments, such as planetary atmospheres. The GEISA (Management and Study of Atmospheric Spectroscopic Information) and CDMS (Cologne Database for Molecular Spectroscopy) databases expanded upon the aims of HITRAN, with CDMS in particular focusing on molecules thought to be of interest in the interstellar or circumstellar mediums and in planetary atmospheres. These databases obtain their data from a number of reliable sources. While in many cases high accuracy data may be guaranteed, the main disadvantage of these databases is lack of coverage. Experimental studies are often limited to specific spectral regions, and although the combination of several observations may slightly improve the overall description across the spectrum, the lack of data is still a common occurrence. The matter of self-consistency (or rather lack thereof) can also become highlighted when using data from multiple sources. As well as lack of coverage in line position, data obtained from experimental observation is often limited in the temperature domain, rendering the data only useful for modelling at room-temperature and below; for example, both HITRAN and GEISA specify spectroscopic parameters at room-temperature [75, 124]. At higher temperatures the molecular spectra become increasingly complicated to analyse, therefore making the extraction of spectroscopic data impossible; performing such experiments becomes ever more difficult as temperatures increase. The high-temperature version of HITRAN, HITEMP [128] provides appropriate data for modelling at elevated temperatures, but is limited to five molecular species. Spectral assignment may also be hindered by other practical issues, as described in the previous section. As such, the data provided by these databases are

often insufficient for modelling high temperature environments.

1.4 The ExoMol Project

The work presented in this thesis forms part of a larger endeavour to produce an extensive spectroscopic database for a number of astronomically important molecular species. The main focus of the project, ExoMol [129], is in the production of high-temperature 'line lists' of molecules for the use of characterising the atmospheres of exoplanets and cool stars. This involves the amalgamation of both reliable experimental and *ab initio* data which may be used in a number of different applications for spectral characterisation and simulation. The data provided are also relevant across a larger range of disciplines, for example in the atmospheric understanding of solar system bodies, and in industrial-based chemical processes described previously. The 'basic' format of ExoMol data is in the listing of the physically accessible energy states of a molecular system, and the Einstein A coefficients corresponding to transitions between them.

The spectra of hot molecules contain many more transitions than those at cooler temperatures, and thus the modelling of high-temperature environments where molecular absorbers dominate is difficult. This is where pre-existing molecular databases falter; their limited coverage in the spectral and temperature domains renders their data unsuitable for accurately modelling the environments which have a high density of molecular absorption lines. Even at lower temperatures their spectral range can be limited. The ExoMol project strives to alleviate these shortcomings from a theoretical perspective, by making use of a number of *ab initio* computational methods (often supplemented or refined by experimental observation) in order to produce spectroscopic data which may be suitable for the modelling of such high-temperature, molecule-rich atmospheres. The advantage of a 'first principles' approach is that full coverage of line position and temperature may be achieved while maintaining self-consistency; the high-density of states and transitions necessary for high-temperature modelling is readily computed from the theoretical perspective. However, the main disadvantage in these efforts to produce high accuracy data is in computational expense. Figure 1.1 shows a comparison of intensity data for phosphine at room-temperature between data provided by ExoMol and by CDMS and HITRAN (2008 edition), where it is clearly seen that

the coverage of data in existing databases is limited even at this low temperature; the 2012 edition of HITRAN [75] was revised based on this computed line list.

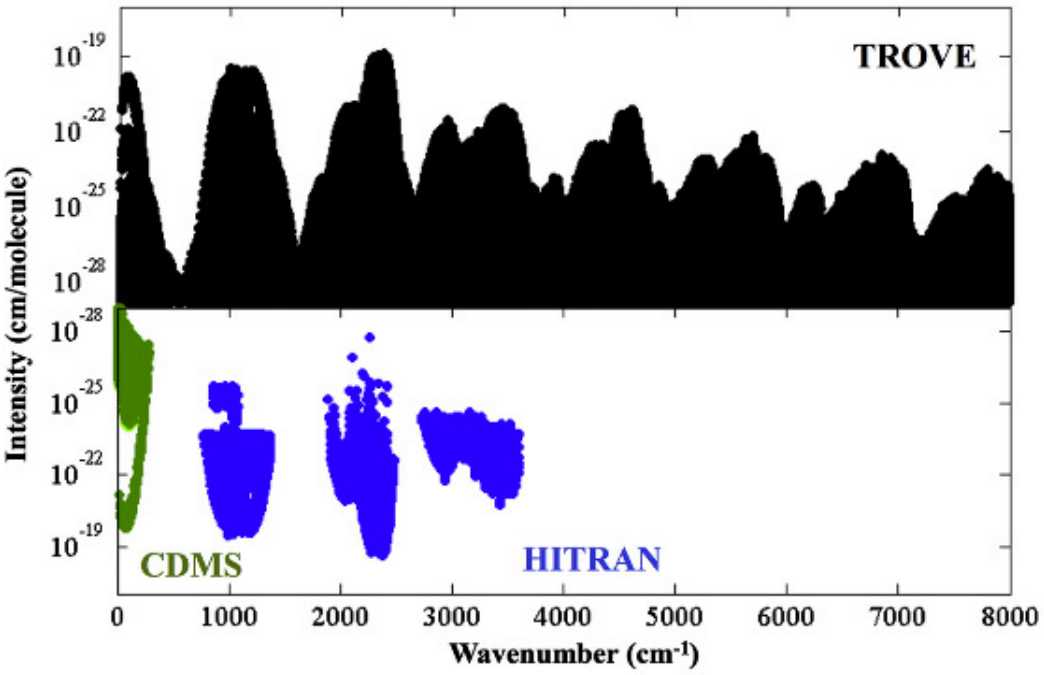

Figure 1.1: Room-temperature line list for PH_3 [130] (above) computed using the TROVE program [131], and corresponding data available in CDMS [76] and HITRAN [75] (below).

ExoMol makes use of multiple high-performance computing facilities in order to meet the high computational demands of hot line list calculations. Due to the difficulty of the task, line list calculations are usually approached on a molecule-by-molecule basis, divided between the ExoMol team and computed by appropriate *ab initio* methods. The convergence of various skills and methodologies within the ExoMol group has allowed for the production of a vast array of data for different molecular species (*www.exomol.com*), and provides an excellent resource for the calculation of spectroscopic data for SO_2 and SO_3, of which this thesis documents in detail.

Line lists already exist for a number of important molecules, including water [132], ammonia [133], methane [134] and phosphine [135], and have already been put to use. For example, the BT2 water line list by Barber *et al.* [132], which contains around 500 million transitions, has been used to identify water in comets, sunspots and a brown dwarf, and for the detection of water in exoplanet HD189733b [136]. The database is continuously updated and extended, and in the future aims to provide additional spectroscopic parameters, including partition functions, cross sections and

broadening parameters [137].

1.5 Thesis Structure

The work presented in this thesis consists of a number of self-contained projects pertaining to theoretical spectroscopic studies of SO_2 and SO_3.

Chapter 2 provides a detailed discussion of the fundamental, theoretical principles upon which all calculations in this work are based. The computational schemes used in the production of spectroscopic data are outlined within the context of both SO_2 and SO_3, including a description of quantum chemistry solutions to the electronic structure problem, and the bespoke algorithms written in order to solve for the nuclear motion of the molecules in question.

Chapter 3 focuses on the work carried out in order to produce a hot line list for $^{32}S^{16}O_2$, via the use of previously computed (and refined) potential energy and dipole moment surfaces (PES and DMS) [35], in conjunction with the DVR3D method [138]. Results are analysed, and compared to previous studies, where possible.

Chapter 4 describes the first portion of a two-part study into the development of a spectroscopically accurate, hot $^{32}S^{16}O_3$ line list, using the TROVE program [131]. Serving as a precursor to the main calculation, the work involves the computation of a room-temperature line list based on a purely *ab initio* PES and DMS, of which the accuracy is scrutinised based on a comparison to available experimental data.

Chapter 5 considers the rotational behaviour of $^{32}S^{16}O_3$ from a theoretical perspective, including both an analysis of the room-temperature rotational (microwave) spectrum and the observation of energy level cluster formation for highly excited rotational states, of which explanation is provided based on the underlying dynamics of the rotating molecule.

Chapter 6 continues from the work of Chapter 4; the *ab initio* PES is spectroscopically refined based on available experimental data, and calculations are extended to higher rovibrational excitations in order to produce a hot line list for $^{32}S^{16}O_3$. The improvement upon the initial line list is analysed in detail.

Finally, Chapter 7 provides some general discussions and conclusions, as well as some considerations for future work.

Chapter 2

Theoretical & Computational Background

The theoretical study of molecular spectra involves a first-principles, quantum mechanical understanding of the behaviour of the electrons and nuclei. The development and improvement of computational tools has allowed for the solution to quantum many-body problems with ever increasing accuracy, and the large library of algorithms is under constant expansion and refinement. Most computational methods pertaining to electronic and nuclear motion can all be attributed to a few common principles, though vary greatly with respect to the mathematical tools used in approaching the problem at hand. This chapter highlights some important principles and methods used to approach both the electronic and nuclear structure problems of small, polyatomic molecules, and how these ideas may be exploited in to understand rovibrational spectra. A comprehensive review of theoretical methods used in the production of rovibrational spectroscopic data for small molecules is given by Lodi and Tennyson [139]. A general overview of those used in the current work is presented in this chapter, within the context of the SO_2 and SO_3 molecules.

2.1 The Molecular Hamiltonian

The mathematical basis upon which quantum mechanics is formed describes the observable values of a quantum system as being represented by vectors in a Hilbert space. In theoretical molecular spectroscopy and quantum chemistry it is of interest to obtain the values representing the total energy of the particular molecular system being studied. For such a system it is possible to construct the molecular Hamiltonian, which

contains the sum of the kinetic and potential energies of Coulomb interactions of all constituent particles, namely the electrons and nuclei, as a function of their spatial distributions:

$$\begin{aligned}\hat{H} = &-\sum_{i=1}^{n}\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{A=1}^{N}\frac{\hbar^2}{2m_A}\nabla_A^2 + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{e^2}{4\pi\varepsilon_0|r_i - r_j|} \\ &-\sum_{i=1}^{n}\sum_{A=1}^{N}\frac{Z_A e^2}{4\pi\varepsilon_0|r_i - R_A|} + \sum_{A=1}^{N}\sum_{B>A}^{N}\frac{Z_A Z_B}{4\pi\varepsilon_0|R_A - R_B|}\end{aligned} \tag{2.1}$$

Here, m_x represents the mass of particle x, e is the electron charge, Z_N is the atomic number of nucleus N, and r_x/R_x represents the coordinate vector of particle x. The constants $\hbar$ and ε_0 refer to the reduced Planck's constant and permittivity of the vacuum, respectively, and ∇_x^2 is the Laplace operator for particle x. The first two terms represent the kinetic energy of the electrons and nuclei, respectively, the third the inter-electron potential energy, the fourth the electron-nuclei potential energy, and finally the fifth term represents the potential energy of Coulombic nuclei-nuclei repulsions. The quantum molecular Hamiltonian takes the form of an operator who's spectrum is the set of possible outcomes for the value of the total energy. One of the principle tasks of molecular physics and quantum chemistry is to obtain the eigenfunctions and eigenvalues of this Hamiltonian for various molecular systems, i.e. by solving the time-independent Schröodinger Equation:

$$\hat{H}\Psi = E\Psi, \tag{2.2}$$

where $\hat{H}$ is the Hamiltonian operator, E is the total energy, and Ψ is an eigenfunction which describes the spatial distribution of the system's constituent particles over all space, otherwise known as the wavefunction. These equations are non-relativistic, and describe the stationary states of a molecular system who's constituent particles are moving at low speeds.

Attacking this problem as presented is a formidable task. This is a many-body problem in which the electronic motion is coupled to the motion of the nuclei and therefore, beyond the solution of relatively small molecular systems, cannot be approached without the implementation of various approximations. The primary concern in the field of quantum chemistry is to obtain the most efficient method of achieving the best solutions by making use of such approximations, the most important of which is the

Born-Oppenheimer approximation, which will be described in the following section.

2.2 The Born-Oppenheimer Approximation

The wavefunction Ψ in Equation 2.2 is a function of electronic and nuclear coordinates describing a many-body problem, and therefore for many molecular systems of interest is extremely difficult to obtain analytically. In an attempt to simplify the problem, it is possible to introduce a decoupling of electronic and nuclear motion. Essentially, it reduces the problem to one whereby the motion of the electrons no longer depends upon the motion of the nuclei, but only on the nuclear positions. The manifestation of this is a separation of the molecular wavefunction into nuclear and electronic components:

$$\Psi_{Mol}(\mathbf{r},\mathbf{R}) = \psi_{Nuc}(\mathbf{R}) \times \psi_{elec}(\mathbf{r};\mathbf{R}) \tag{2.3}$$

Both $\mathbf{r}$ and $\mathbf{R}$ are position vectors representing all electronic and nuclear coordinates, respectively. The electronic wavefunction ψ_{elec} has a parametric dependence on the instantaneous position of the nuclei, $\mathbf{R}$. This decoupling of the nuclear and electronic motions is the core principle of the Born-Oppenheimer approximation [140]. The justification of this approximation is usually explained by the large ratio of the nuclear and electron masses; both the nuclei and electrons experience coulombic forces of similar magnitude yet, since the electrons are so much lighter they move much faster relative to the nuclei. Therefore the assumption is that the electrons redistribute themselves instantaneously upon any nuclear motion.

In this picture, the kinetic energy of the nuclei is neglected in solving for the total electronic energy, i.e. the nuclei are 'clamped'. With this taken into consideration, the remaining components of Eq. 2.1 can define the clamped nucleus Hamiltonian:

$$\begin{aligned}\hat{H}_{cn} =& -\sum_{i=1}^{n}\frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0}\Bigg[-\sum_{i=1}^{n}\sum_{A=1}^{N}\frac{Z_A}{4|r_i-R_A|} \\ &+\sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{|r_i-r_j|}\Bigg] + \sum_{A=1}^{N}\sum_{B>A}^{N}\frac{Z_A Z_B}{4\pi\varepsilon_0|R_A-R_B|},\end{aligned} \tag{2.4}$$

The resulting Schröodinger equation takes the form

$$\hat{H}_{cn}(\mathbf{r};\mathbf{R})\psi_{elec}(\mathbf{r};\mathbf{R}) = E_{cn}(\mathbf{R})\psi_{elec}(\mathbf{r};\mathbf{R}), \tag{2.5}$$

where both $\psi_{elec}(\mathbf{r};\mathbf{R})$ and $E_{cn}(\mathbf{R})$ depend on the nuclear coordinates $\mathbf{R}$. Using this framework, the electronic energy problem is solved for various configurations of fixed nuclear coordinates. This gives rise to the well-known concept in quantum chemistry of a potential energy surface (PES), a grid of electronic energies corresponding to each nuclear arrangement.

This is the first part of a two-step process. The so called electronic structure problem is solved for a number of nuclear geometries in order to construct the PES. It is possible for multiple electronic configurations to exist for a particular fixed nuclear geometry, depending on the particular system being described, and this would lead to a number of potential surfaces existing which correspond to each electronic arrangement. For the remainder of this work, the only concern is with the ground electronic state, which gives rise to a single PES for each molecule of interest.

The second step makes use of the constructed PES in solving the for the motion of the nuclei, with the nuclear Schröodinger equation taking the form

$$\left[-\sum_{A=1}^{N}\frac{\hbar^2}{2m_A}\nabla_A^2 + E(\mathbf{R})\right]\psi_{nuc}(\mathbf{R}) = E_{nuc}\psi_{nuc}(\mathbf{R}), \tag{2.6}$$

where the previously neglected nuclear kinetic energy term has been reincorporated.

The Born-Oppenheimer approximation described here is the at the root of all calculations described in this work, where various methods are used to solve the two step process. The following section describes the general approach to solving the first step, the problem of electronic motion.

2.3 The Electronic Motion Problem

Despite the simplification that the Born-Oppenheimer approximation strives to achieve, finding exact solutions to the clamped nucleus Hamiltonian is still analytically impossible for most molecular systems of interest. Instead, numerical procedures are implemented in order to solve these eigenvalue problems *ab initio*. This section briefly outlines the principles of electronic structure theory, and mentions various quantum chemistry methods that are used in the calculations for constructing potential surfaces, particularly the coupled-cluster Theory which has been used in the construction of the PESs used in this thesis.

2.3.1 The Hartree-Fock Method & Post Hartree-Fock Methods

The fundamental approach underlying most *ab initio* electronic structure calculations lies in the Hartree-Fock method [141–143]. The starting point for this particular scheme is an anzatz for the electronic wavefunction, taken to be a product of one-electron functions, or 'orbitals', which depend on the position vector, $\mathbf{r_i}$ of electron i. The existence of spin adds an extra degree of freedom alongside the spatial coordinates of the electrons, so spin-orbitals are generally used, denoted by χ :

$$\psi_{elec} = \chi_1(\mathbf{r}_1)\chi_2(\mathbf{r}_2)...\chi_N(\mathbf{r}_N) \tag{2.7}$$

A direct product of one-electron functions however does not satisfy the Pauli principle in that the wavefunction is not antisymmetric with respect to the interchange of two identical electrons (which are fermions). Instead, a linear combination of such products can be used to account for this necessity. This can be represented as a Slater determinant:

$$\psi_{elec} = \psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{r}_1) & \chi_2(\mathbf{r}_1) & \cdots & \chi_N(\mathbf{r}_1) \\ \chi_1(\mathbf{r}_2) & \chi_2(\mathbf{r}_2) & \cdots & \chi_N(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{r}_N) & \chi_2(\mathbf{r}_N) & \cdots & \chi_N(\mathbf{r}_N) \end{vmatrix} \tag{2.8}$$

In order to generate the best approximation for the electronic wavefunction it is possible to make use of the variational principle, which states that the energy of a given trial wavefunction is greater than or equal to the value of the exact energy; in the case of such an equality, the trail wavefunction is the exact wavefunction. By use of this principle it is possible to vary the parameters of the initial wavefunction anzatz to minimise the energy.

In this single-Slater determinant picture, it is assumed that each electron interacts with an average charge distribution which is caused by the remaining electrons. This is incorrect and can lead to a poor description of the electronic behaviour - it does not correctly take into account the so-called Coulomb correlation, an effect arising from the fact that the Hartree-Fock model treats instantaneous Coulumbic repulsions between electrons in a simplified way by having individual electrons interact with an average

electron charge cloud. This leads to an energy that is different from the exact energy. The difference between the exact energy and the Hartree-Fock energy is known as the correlation energy [139]:

$$E_{corr} = E_{exact} - E_{HF} \tag{2.9}$$

This is one of the primary shortcomings of the Hartee-Fock method. Although it allows for some correlation between electrons of identical spin, it neglects any between those of opposite spin. The Hartree-Fock method can be complemented by additional schemes which aim to account for the residual correlation energy.

The aim of post Hartree-Fock methods is to improve upon the accuracy of the HF calculation. This does however come at a price, that of computational expense. There are a variety of methods available aiming to achieve the desired accuracy, and each have their own computational costs; generally speaking, the higher the accuracy of the result required, the more computationally demanding.

The most well-known post Hartree-Fock methods include the configuration interaction (CI) method, Møller-Plesset perturbation theory (MP) and coupled-cluster (CC) theory [139, 144–146]. In the case of CI, the solution to the Hartree-Fock method is assumed to be a good approximation, and the 'true' wavefunction is expanded as a weighted, linear combination of Slater determinants, including the Hartree-Fock ground-state determinant, as well as determinants representing excitations to unoccupied (virtual) orbitals, which contribute less and less to the overall wavefunction with increasing excitation [139]. This method aims to account for the pair correlation of electrons, and optimises the appropriate weights in a minimisation scheme. Due to computational difficulty the wavefunction expansion is often truncated to include single excitations, single and double excitations, etc. from the reference Hartree-Fock wavefunction, and methods are referred to by acronyms such as CIS, CISD, etc. These all offer varying degrees of quality in final results, with inclusion of higher order excitations providing greater accuracy.

The coupled-cluster approach is similar in that it relies on a good approximation in the Hartree-Fock calculation, and improves upon the CI method, most notably by considering the problem of size-consistency, which requires the total energy of a combined system to be equal to the individual energies of the system's constituents separated

by a large enough distance where their interaction is zero [139]. The coupled-cluster wavefunction is given as:

$$\psi_{CC} = e^{\hat{T}} \psi_0, \tag{2.10}$$

where ψ_0 is the a reference wavefunction, i.e. a Hartree-Fock wavefunction, and $\hat{T}$ is the cluster operator, defined as:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ... \tag{2.11}$$

Here, the $\hat{T}_i$ are the cluster operators for single, double, triple etc. excitations. The inclusion of higher order excitations results in a convergence to the exact value within a one-electron basis, i.e. full correlation [147]. In practice, similar to the CI method, $\hat{T}$ is truncated due to the increasing computational demands of higher excitations. As such, the naming convention of these techniques is similar to that of CI and is defined by the order of truncation, with CCS referring to a coupled-cluster computation using only single excitations, CCSD for singles and doubles, CCSDT for singles, doubles and triples, etc. There are also examples of coupled-cluster approaches introducing approximate treatment for higher excitation, usually denoted by parentheses; e.g. CCSD(T) refers to a coupled-cluster calculation for single and double excitations, with perturbative treatment of triple excitations. This method in particular is well-known to provide good accuracy while minimising computational costs, and is often referred to as the 'gold standard' [139]. This approach is used for all calculations discussed in this thesis. The coupled-cluster method is reviewed in greater detail in Ref. [147].

The accuracy to with which results are achieved is also dependent on the choice of the basis set used to construct the one-electron functions in the Hartree-Fock Slater determinant [144]. Typically these are composed as linear combinations of atomic orbitals, with the principle aim of describing the behaviour of an electron around the atomic nucleus as well as possible. The 'Slater type orbitals' (STOs) generally give a good description of such behaviour, however they are computationally difficult to handle. For this reason they are more often approximated by contracted linear combinations of 'Gaussian type orbitals' (GTOs), in an effort to emulate the behaviour of STOs.

A minimal basis set is one which contains the minimum number of basis functions

to describe each atom, the most common of which is the STO-nG basis set [148]. Here, the n represents the number of 'primitive' Gaussian functions used in the construction of an individual basis function; the linear combination is referred to as a 'contraction'. The resulting STO-nG is subsequently used as an atomic orbital basis function. This representation can be expanded upon to better describe the nature of molecular bonding by providing an improved description of the valence electron orbitals. Such a description is known as a 'split-valence' basis set [149]. In this instance, the core orbital remains described by a single contracted combination of Gaussian functions, however the basis set used to describe the valence electrons are comprised of a linear combination of contracted Gaussians. A basis set described in this way is denoted as X-YZG, where X represents the number of primitive Gaussians used in describing the core orbital, and the Y and Z represent the number of primitives for each Gaussian contraction used in constructing the valence orbital. It is possible to add more Gaussian contractions to the valence orbital description, therefore increasing from what is known as a 'split-valence double-zeta' basis set (seen here) to 'triple-zeta', 'quadruple-zeta', etc.

Further additions to these basis set descriptions can be added, including the use of polarisation functions, in which contracted Gaussian functions describing higher excitations of angular momentum for the valence orbitals may be included to enhance the 'flexibility' of the basis set. Another set of similar enhancements are the inclusion of diffuse functions, in which the contracted Gaussians describe the behaviour of electrons far from the nucleus. Each of these can help provide a less restrictive description of the orbital behaviour. The former are usually denoted with an asterisk in the standard Pople naming convention, whereas the latter are described with a '+'.

The Pople basis sets described here are not well suited to the correlated, post-Hartree Fock computations discussed previously. Dunning [150] provided 'correlation-consistent' basis sets optimised to work with such methods, whose descriptions follow a similar pattern to the Pople basis sets. Namely, double, triple, quadruple-zeta, etc are specified along with any additional properties, such as polarisation functions and diffuse augmentations. For example, the basis set named 'aug-cc-pVTZ' is augmented with diffuse functions, is correlation-consistent, polarised-valence, and is a triple-zeta combination.

There are several basis sets available, each suited to help solve the particular prob-

lem described. More information on basis sets and electronic structure methods in general is given in Ref. [139], and references within. Figure 2.1 [151] shows a plot outlining the quality of *ab initio* electronic structure methods and basis sets with respect to obtaining exact solutions; as can bee seen, the extension of both basis set size and electron correlation consideration is required to achieve experimental accuracy.

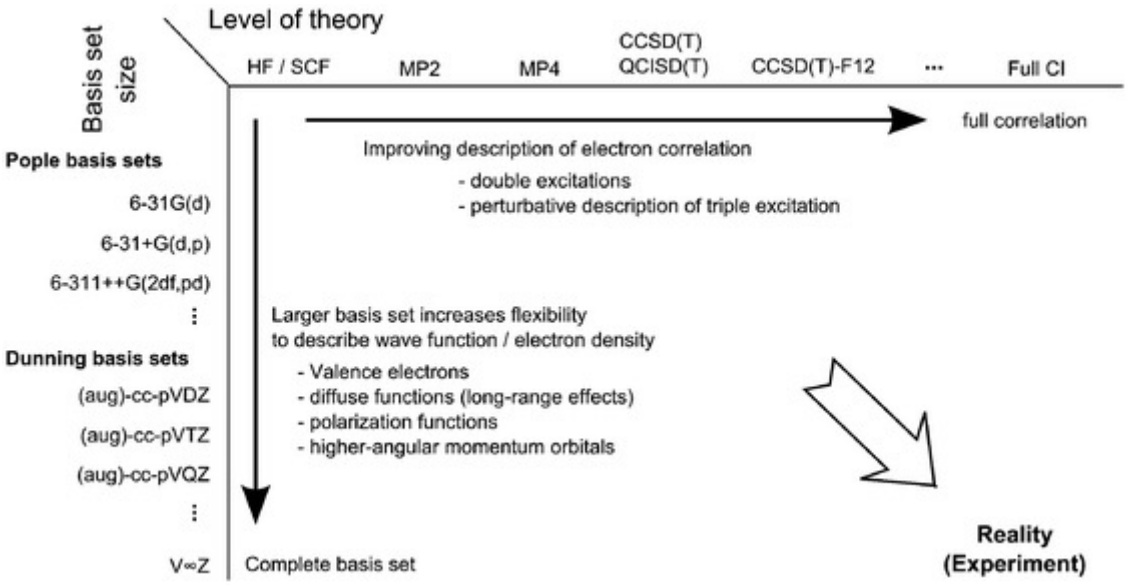

Figure 2.1: Hierarchy of electronic structure methods and basis set size. Figure taken from Ref. [151].

2.4 The Nuclear Motion Problem: A Precursor

Once the electronic problem is dealt with it is necessary to solve for the motion of the nuclei, i.e. by solving Eq. 2.6. Various methods have been implemented to achieve this, many of which work by receiving an externally computed potential energy surface as an input, with which subsequent calculations are made. The work in this thesis uses two different methods to compute the rovibrational wavefunctions, energies and dipole transition moments for the two molecular systems of interest. Both methods involve the construction of a Hamiltonian matrix as a sum of kinetic and potential energy operators. Diagonalization of these matrices in terms of suitable basis set functions gives the necessary eigenfunction and eigenenergies desired.

In the next three sections these methods will be addressed. First, a general discussion of the fundamental approaches common to both methods is given. In the following section an overview of the DVR3D method is given. DVR3D specialises in the compu-

tation of rovibrational energies and eigenfunctions for triatomic systems; the procedure is outlined and then discussed in the context of SO_2. The final section regarding the nuclear motion problem deals with the TROVE method. TROVE employs a more general computational scheme for arbitrary polyatomic molecules, and has been adapted to work with tetratomic systems. This is the technique used in the calculations of rovibrational energies for SO_3.

2.4.1 Overview of the Problem

The aim is to solve the Nuclear Schröodinger equation of a polyatomic molecule which, in Cartesian coordinates, can be expressed as

$$\left[-\frac{\hbar^2}{2}\sum_{A=1}^{N}\frac{\nabla_A^2}{m_A}+V\right]\psi_{TRV}=E_{TRV}\,\psi_{TRV}, \tag{2.12}$$

where each nucleus A has mass m_A and coordinates (R_{AX},R_{AY},R_{AZ}) in the laboratory-fixed axis system. V is the potential surface obtained from solution to electronic structure calculations outlined in the previous section, and the ψ_{TRV} and E_{TRV} are the eigenfunctions and eigen-energies, respectively. The subscript TRV denotes that the solutions are for the energies associated with translational, rotational and vibrational motion.

For a molecule with N atoms, there are a total of $3N$ degrees of freedom in the laboratory fixed frame associated with these motions. In practice, the Schröodinger Equation for the N-atomic molecule in this Cartesian form is not easy to solve. It is possible to introduce more suitable coordinates which take into account the fact that the translational motion can be completely separated from the rotational and vibrational motion, and that the remaining motions can also (to a lesser extent) be decoupled [131, 152, 153], thereby simplifying the construction and eventual solution of the Schröodinger Equation.

To this end, a transformation to a 'space-fixed' axis system is introduced, where the origin lies at the centre of mass of the molecular system, thus removing the centre of mass motion entirely. This leaves a total of $3N$ - 3 degrees of freedom associated with rotation and vibration. In reality it is not possible to completely decouple these motions, but to a good first approximation the dimensionality of the problem may be

further reduced by introducing the 'molecule-fixed' axis, which essentially rotates with the molecule; in this reference frame the only motion is from vibrations of the molecule, where the removal of the rotational degrees of freedom leaves $3N$ - 6 vibrational degrees (or $3N$ - 5 for linear molecules).

This, together with the separation of the translational motion, allows for the 'rovibrational' eigenfunction to be defined in terms of suitable coordinates which describe both the rotational and vibrational motions:

$$\Psi_{RoVib} = \psi_{rot}(\theta, \phi, \chi)\psi_{vib}(q), \tag{2.13}$$

where θ, ϕ and χ are the Euler angles [154], and q are a set of internal coordinates. In order to define the orientation of the molecule with respect to the space-fixed frame the Euler angles are introduced as parameters of an orthogonal rotation from the molecule-fixed frame to the space-fixed frame, i.e. they describe the rigid rotations of the molecule [153]. It is therefore useful to use standard rigid rotor functions in terms of these angles in order to describe this motion. The remaining $3N$ - 6 internal coordinates and corresponding basis functions must be chosen to describe the vibrational motions. Examples of such functions are standard harmonic or spherical functions, in terms of bond length or bond displacement coordinates, and inter-bond angles.

The possible transformations to the molecule-fixed axis system are described more rigorously in Refs. [131,153]. Despite the differences in methodology, the common aim of both DVR3D and TROVE is to construct a Hamiltonian for the molecule in question in terms of appropriate internal coordinates and basis functions. Both the DVR3D and TROVE program suites operate on the assumption of this coordinate transformation to the molecule-fixed frame, and each allow for the choice of suitable coordinates and basis functions for describing the vibrational motions. As a result, the potential energy surface, V, which is supplied as an external function, must be given in terms of the internal coordinates used to describe the problem.

Thus, a matrix representation of the Hamiltonian operator is constructed in accordance with the chosen coordinates, and diagonalised in terms of a suitable basis set comprised of functions describing the rotational and vibrational motions.

2.4.2 The Finite Basis Representation

The aim is to diagonalize the Hamiltonian matrix in terms of a basis set which consists of a product of 1-dimensional basis functions, describing individual bond stretches, inter-bond bends, and rotation. Historically, the usual approach to representing a basis set involves a Finite Basis Representation (FBR), which is a truncation of an infinite Variational Basis Set (VBR). In the VBR, wavefunctions are expressed as a superposition of orthonormal basis functions:

$$\psi = \sum_{i=1}^{\infty} c_i \phi_i. \tag{2.14}$$

In the variational approach, the coefficients c_i are optimised to achieve a minimisation in the eigenvalue associated with ψ. Eigenvalues of the Hamiltonian matrix are found by diagonalizing the matrix represented in the chosen VBR. In practise, an FBR is used by truncating this sum to retain N basis functions, ϕ_i. This truncation introduces an approximation to the minimised result for the eigenvalues. As the basis approaches completeness the eigenvalues of the Hamiltonian matrix converge to their true values, however due to the increased computational demands of a larger basis set, there must be a trade-off between accuracy and computing time, which is why the truncation is introduced.

2.4.3 Symmetry

It is often practical to take advantage of the symmetrical properties inherent in many molecules to further simplify the computation of rovibrational wavefunctions. One method is to classify the molecule according to a Molecular Symmetry (MS) group, under the mathematical framework of group theory [154]. This classification is based on the geometry of the molecule. A so-called 'point group' contains various symmetry operations across elements (axes, planes, etc.); in the context of a molecule these symmetry operations may be characterised as nuclear permutations and inversions. The labelling of wavefunctions via the MS group is based on the behaviour of the wavefunction when subject to the operations of the group.

A Molecular wavefunction describing a rovibrational state of a system can be labelled by an irreducible representation (or 'symmetry species') of the MS group, de-

pending on how it transforms under the various operations of the group, and this is useful for a number of reasons. Firstly, the classification of functions by symmetry species allows for the determination of Pauli Principle-permitted wavefunctions. From a computational perspective this helps eliminate the calculation of unnecessary wavefunctions which describe unphysical molecular states.

Secondly, the symmetry labelling of molecular functions introduces simplification to calculations involving integrals; in particular, it is of interest to determine the conditions which integrals of the form $\langle \psi_i | \hat{O} | \psi_j \rangle$ must satisfy in order to be non-zero, where $\hat{O}$ is an operator. Integrals of this form appear in the construction of matrix elements in the Hamiltonian matrices, and the simulation of intensities via the computation of transition strengths between states and their corresponding rovibrational wavefunctions.

It follows that there is a dependency upon the symmetry properties of the individual functions involved in the integrand, ψ_i, $\hat{O}$, and ψ_j. If these functions are from spaces whose bases generate the irreducible representations Γ_i, $\Gamma_{\hat{O}}$ and Γ_j, respectively, and if $\Gamma_{\hat{O}} \otimes \Gamma_j$ does not contain Γ_i, then [154]

$$\langle \psi_i | \hat{O} | \psi_j \rangle = 0. \tag{2.15}$$

Therefore, with prior knowledge of the symmetry behaviour of ψ_i, $\hat{O}$, and ψ_j, it is possible to avoid unnecessary calculations of integrals that would ultimately equate to zero. An additional consequence is that transition selection rules are imposed by this rule; these are discussed Section 2.7.

2.5 The Nuclear Motion Problem: SO_2 & DVR3D

The nuclear motion rotation-vibration wavefunctions required for computing dipole transition moments for the focus of the current work on SO_2 are calculated using the DVR3D program suite [138]. This set of program modules has been built specifically for triatomic systems, based on the Discrete Variable Representation (DVR). The DVR was first used by Harris *et al.* [155] and was applied to some simple one-dimensional quantum mechanical problems. Its potential was not discovered until much later when Light *et al* [156, 157] devised a procedure for calculating nuclear motion states of molecules, the principles of which define the workings of DVR3D.

The DVR3D suite allows for the calculation of ro-vibrational wavefunctions and spectra using either Radau or Jacobi coordinates. There are a number of advantages of using the DVR technique above other widely used variational techniques. These, along with a background of the DVR method are discussed.

2.5.1 The SO_2 Molecule

SO_2 is a bent, asymmetric top molecule characterised by $\mathscr{C}_{2v}$ symmetry, with the sulphur atom placed on the C_2 symmetry axis. SO_2 has three fundamental vibrational modes (see Fig. 2.2), which are the ν_1 symmetric stretch, the ν_2 symmetric bend and the ν_3 asymmetric stretch, functions of which transform as A_1, A_1 and B_2 under $\mathscr{C}_{2v}$, respectively. The Bosonic nature of the spin-0 Oxygen atoms means that the total rovibrational wavefunction must be symmetric with respect to the permutation of the Oxygen nuclei, as per the Pauli Principle. From the $\mathscr{C}_{2v}$ character table it can be shown that this condition is satisfied only for the A_1 and A_2 symmetry species [154].

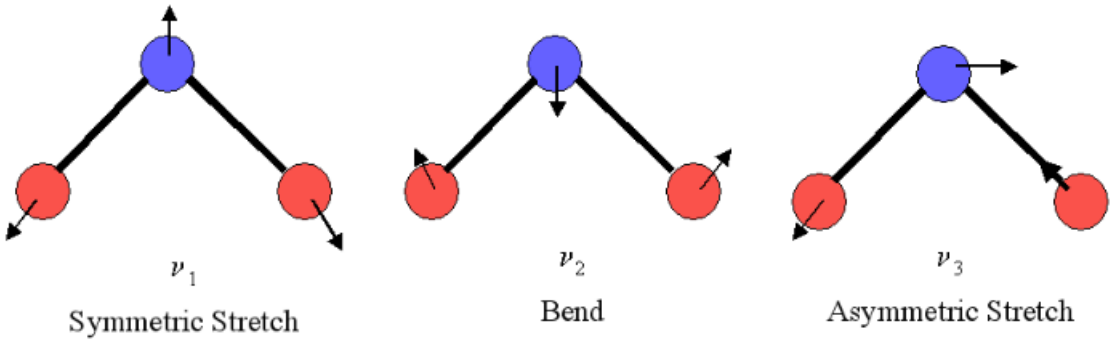

Figure 2.2: Fundamental vibrations of SO_2.

The DVR3D suite deals with the labelling of rovibrational states via vibrational 'symmetry blocks' defined by the vibrational basis symmetry, q, and rotational parity, p, which are either even (e) or odd (o) [138] (see Subsection 2.5.4). Thus, there are four possible symmetry block combinations qp for vibrational and rotational basis functions: ee, eo, oe, oo. These blocks can be related to the irreducible representations of $\mathscr{C}_{2v}$ [154], but the relation is dependent on whether the rotational angular moment quantum number J is even or odd. This is summarised in Table 2.1.

Since the Pauli Principle physically excludes states with B-type symmetry for SO_2, only two symmetry block combinations are required for each J, which simplifies the calculations by disregarding half of all the possible rovibrational wavefunctions. The

Table 2.1: Symmetry block combinations within the DVR3D suite, and their relations to the irreducible representations of $\mathscr{C}_{2v}$ [154].

	***J* even**				***J* odd**			
q	e	e	o	o	e	e	o	o
p	e	o	e	o	e	o	e	o
$\Gamma(\mathscr{C}_{2v})$	A_1	B_1	B_2	A_2	B_1	A_1	A_2	B_2

calculation of these is defined by certain input parameters to the DVR3D code which are discussed in Chapter 3. The method for their calculation is now discussed.

2.5.2 The Coordinate System, Orientation Within the Molecule-fixed Frame, & the Finite Basis Functions

There are two choices of orthogonal coordinate systems implemented in the DVR3D program suite - Jacobi and Radau coordinates. The Jacobi, or 'scattering' coordinates (r_1, r_2, θ) are well suited for describing collision systems between a single particle and a diatom [158]. In this instance, r_1 represents the bond length of the general diatomic BC, and r_2 the distance of the weakly bound atom A from the centre of mass of BC, where θ represents the angle between these two vectors.

For the purposes of the current work on SO_2, the Radau coordinates are used instead. Here r_1 and r_2 represent the distances of atoms 1 (O_1) and 2 (O_2), respectively, from a 'canonical point' P. This point satisfies the condition $\overline{MB} \times \overline{CB} = \overline{PB}^2$, where B is the centre of mass of atoms 1 and 2, C is the centre of mass of the triatomic system, and M is the distance of atom 3 (S) from point B [159]. In the case of a heavy central atom 3 (relative to the other atoms), the point P has little variation, and usually coincides with the centre of mass of the third atom [158]. The angle between r_1 and r_2 is represented by θ. Figure 2.3 shows a graphical representation of the Radau coordinate system.

DVR3D specifies a set of generalised internal coordinates, g_1 and g_2, where [138, 160]

$$g_1 = \frac{A_3 - P}{A_3 - A_2}, \qquad g_2 = \frac{A_3 - R}{A_3 - A_1}. \tag{2.16}$$

Here P and R specify the centre-of-masses between atom 3 and 2, and 3 and 1, respectively, as in Figure 2.4.

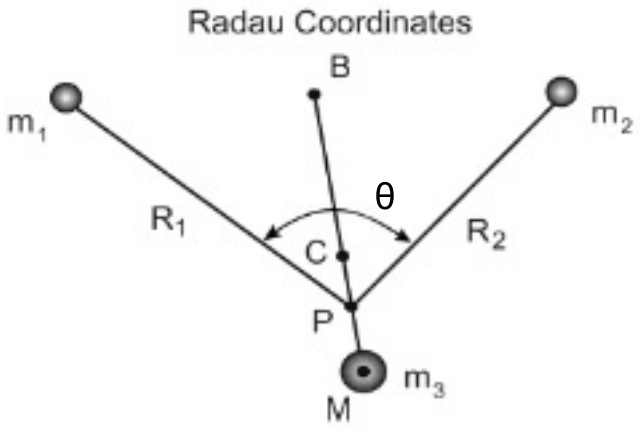

Figure 2.3: The Radau coordinate system [159].

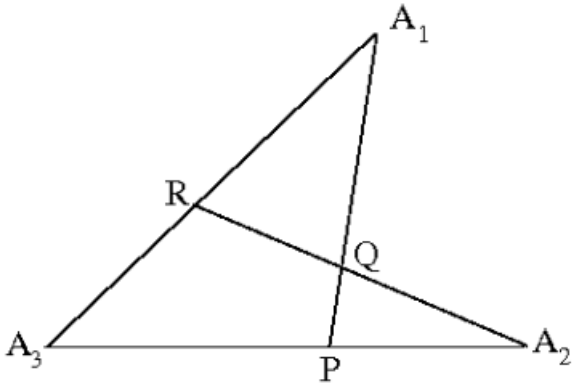

Figure 2.4: The generalised coordinate system. A_i represents atom i [138, 160].

In general, these coordinates can be obtained for any system [160], by using

$$g_1 = 1 - \frac{A}{A+B-AB}, \qquad g_2 = 1 - \frac{A}{1-B+AB}, \tag{2.17}$$

$$A = \left(\frac{m_3}{m_1+m_2+m_3}\right)^{1/2}, \qquad B = \frac{m_2}{m_1+m_2}, \tag{2.18}$$

and the m_i are the masses of atoms 1, 2 and 3. In the case where m_3 is much larger than m_1 and m_2, the Radau coordinates behave as bond-length-bond-angle coordinates.

The choice of the orientation of the molecule, or 'embedding', within the molecule-fixed frame xyz is important for facilitating the solutions to the rotational excitation problem [138]. DVR3D includes four different types of embedding options:

1. r_1 embedding: z-axis parallel to r_1, with x in the plane of the molecule;

2. r_2 embedding: z-axis parallel to r_2, with x in the plane of the molecule;

3. Bisector embedding: x bisects the angle θ, with z in the plane of the molecule;

4. Perpendicular embedding: x bisects the angle θ, with z perpendicular to the plane of the molecule.

The current work makes use of the bisector embedding [138,161], which takes advantage of the symmetric properties of SO_2; the z-axis coincides with the $\mathscr{C}_{2v}$ symmetry axis.

2.5.2.1 The Basis Functions

It is practical to use a set of orthogonal basis functions to describe each of the stretching motions in the r_1 and r_2 coordinates and the bending motion in the θ coordinate. DVR3D uses Laguerre polynomials for the radial basis functions, in the form of either Morse or spherical oscillators; for SO_2 the Morse-type function [153, 162] is used:

$$|n\rangle = \beta^{1/2} N_{n\alpha} exp(-y/2) y^{(\alpha+1)/2} L_n^{\alpha}(y), \qquad y = Aexp[-\beta(r-r_e)], \tag{2.19}$$

where $N_{n\alpha} L_n^{\alpha}(y)$ is a normalised associated Laguerre polynomial [153, 163], and

$$A = \frac{4D_e}{\omega_e}, \qquad \beta = \omega_e(\mu/2D_e)^{1/2}, \qquad \alpha = INT(A), \tag{2.20}$$

Here, μ is the reduced mass, and D_e, ω_e and r_e are parameters in the Morse potential function:

$$V(r) = D_e(1 - exp[-\beta(r-r_e)])^2 - D_e. \tag{2.21}$$

D_e, ω_e and r_e are treated as variational parameters in the DVR3D computation, and their values must be optimized accordingly in order to minimize and converge the basis set. Such convergence tests are outlined in Chapter 3. Since SO_2 is homonuclear, the same basis functions may be used for both the r_1 and r_2 coordinates.

In the case of the bisector embedding that is made use of in the current project, the angular basis functions, $|l\rangle$, are Legendre polynomials in the case of the purely vibrational problem, and associated Legendre polynomials of order k in the case of rotational excitation [138] (see below).

Thus, in the FBR picture, the state $|\psi\rangle$ is expanded in terms of the basis functions

as

$$|\psi\rangle = \sum_{i,j,k} c_{ijk} l_i(\theta) m_j(r_1) n_k(r_2). \tag{2.22}$$

As has been previously stated, the FBR introduces a truncation in the number of basis functions that are retained, therefore the representation of the true wavefunction is compromised, particularly for high lying states. The lack of convergence of the eigenvalues of these states may be improved by the addition of more basis functions, however this becomes more computationally expensive.

2.5.3 The Discrete Variable Representation

The defining feature of DVR3D is the introduction of the Discrete Variable Representation (DVR) [138, 156, 157], which aims to mitigate the problems associated with the FBR. The approach provides an efficient way in constructing the 3D Hamiltonian in the Radau coordinates. The benefit of the DVR is that it provides a simple method for evaluating matrix elements, and a compact way of representing the multidimensional Hamiltonians.

The DVR is obtained as an orthogonal transformation from an equivalent FBR. Assuming that the FBR is comprised of a set of orthogonal polynomials, the transformation is defined in terms of the Gaussian quadrature points and weights of these polynomials; for an FBR consisting of N orthogonal polynomials, there exists an isomorphic DVR at N weighted quadrature points. Essentially, the transformation changes from a representation where the wavefunctions are expressed as amplitudes associated with the polynomial functions to one where they are expressed as amplitudes at the quadrature points.

A 1-dimensional DVR transformation, T, can be represented by the points and weights of the N-point Gaussian quadrature rule associated with the corresponding FBR polynomial in the coordinate of interest (r_1, r_2, θ):

$$T_{\gamma l} = \sqrt{w_\gamma} l(\gamma), \qquad T_{\alpha m} = \sqrt{w_\alpha} m(\alpha), \qquad T_{\beta n} = \sqrt{w_\beta} n(\beta), \tag{2.23}$$

where the l_i, m_j, n_k are in accordance with Eq. 2.22.

Since the FBR is expressed as a direct product of the polynomial functions, a 3D transformation to the DVR can be defined as the product of the 1D transformation

operators:

$$\underline{T} = T_{\alpha m} T_{\beta n} T_{\gamma l}. \tag{2.24}$$

2.5.4 Method

In the following a brief outline of the DVR3D method is presented, with emphasis on the technique for constructing rovibrational wavefunctions for homonuclear triatomics in the bisector embedding [138, 161], used in the current work on SO_2.

For the the purely vibrational $J = 0$ case, the Hamiltonian can be written in terms of the FBR as

$$\begin{aligned}\langle m,n,l|H^{J=0}|m',n',l'\rangle &= \langle m|\hat{h}^{(1)}|m'\rangle\delta_{n,n'}\delta_{l,l'} + \langle n|\hat{h}^{(2)}|n'\rangle\delta_{m,m'}\delta_{l,l'} \\ &+ (\langle m|\hat{g}^{(1)}|m'\rangle\delta_{n,n'} + \langle n|\hat{g}^{(2)}|n'\rangle\delta_{m,m'})l(l+1)\delta_{l,l'} \\ &+ \langle m,n,l|V(r_1,r_2,\theta)|m',n',l'\rangle,\end{aligned} \tag{2.25}$$

where

$$\hat{h}^{(i)} = \frac{-\hbar^2}{2\mu_i}\frac{\partial^2}{\partial r_i^2}, \qquad \hat{g}^{(i)} = \frac{\hbar^2}{2\mu_i r_i^2}, \tag{2.26}$$

and μ_i are appropriate reduced masses in terms of the parameters of Equations 2.17 and 2.18 [138].

Here, the radial basis functions $|m\rangle$ and $|n\rangle$ are the Laguerre-type polynomials discussed in Subsection 2.5.2.1, and the angular functions $|l\rangle$ are Legendre polynomials. Since SO_2 is a symmetric AB_2 type molecule it makes sense to take advantage of the inherent symmetry present by using the same coordinates and basis functions to describe the radial behaviour. The resulting symmetrised radial functions are given as

$$|m,n,q\rangle = \sqrt{2(1+\delta_{m,n})}(|m,n\rangle + (-1)^q|n,m\rangle), \qquad q = 0,1, \tag{2.27}$$

where q denotes the even ($q = 0$) and odd ($q = 1$) symmetry blocks.

DVR3DRJZ is used as a first step in the computation of the full rovibrational wavefunctions. This first step involves the calculation of a Coriolis-uncoupled Hamiltonian, $H^{J,k}$ for $J > 0$. In this picture, it is assumed that the projection of J onto the molecule-fixed z-axis, k, is a good quantum number. Here, k is different from the well-known

asymmetric top quantum numbers K_a and K_c [154], however at equilibrium k corresponds to K_a [164]. The angular basis $|l\rangle$ functions for $J > 0$ are associated Legendre polynomials of order k, and are coupled to symmetrised rotor functions [138], denoted by $|J,k,p\rangle$, where p defines the rotational parity by $(-1)^{J+p}$ ($p = 0$ and $p = 1$ for even and odd parity, respectively).

In the FBR, for the bisector embedding, $H^{J,k}$ is given as [138]

$$\begin{aligned}\langle m,n,l,J,k,p|H^{J,k}|m',n',l',J,k,p\rangle =& \langle m,n,l|H^{J=0}|m',n',l'\rangle \\ &+ (\langle m|\hat{g}^{(1)}|m'\rangle\delta_{n,n'} + \langle n|\hat{g}^{(2)}|n'\rangle\delta_{m,m'}) \\ &\times \Big(\frac{1}{4}\langle lk|\frac{1}{(1-cos\theta)}|l'k\rangle(J(J+1)-3k^2) \\ &+ \frac{1}{8}\delta_{l,l'}(J(J+1)-k^2) \\ &+ \delta_{k,1}\frac{(-1)^{1-p}}{16}\langle lk|\frac{1+cos\theta}{(1-cos\theta)}|l'k\rangle J(J+1)\Big).\end{aligned} \tag{2.28}$$

Here there is an extra term on the diagonal for the case $k = 1$, due to the symmetrisation of the rotational functions [138, 161].

The Hamiltonian of Equation 2.28 may be transformed to the DVR associated with the quadrature schemes of the FBR basis functions via the composite, unitary transformation operator in Eqaution 2.24:

$$\underline{T}^T\langle m,n,l,J,k,p|H^{J,k}|m',n',l',J,k,p\rangle\underline{T}. \tag{2.29}$$

This gives the transformed Hamiltonian written at the DVR points

$$\begin{aligned}H^{J,k}_{\alpha,\alpha',\beta,\beta',\gamma,\gamma'} =& K^{(1)}_{\alpha,\alpha'}\delta_{\beta,\beta'}\delta_{\gamma,\gamma'} + K^{(2)}_{\beta,\beta'}\delta_{\alpha,\alpha'}\delta_{\gamma,\gamma'} + L^{(1)}_{\alpha,\alpha',\gamma,\gamma'}\delta_{\beta,\beta'} + L^{(2)}_{\beta,\beta',\gamma,\gamma'}\delta_{\alpha,\alpha'} \\ &+ \frac{1}{8}(M^{(1)}_{\alpha,\alpha',\beta,\beta'} + M^{(2)}_{\alpha,\alpha',\beta,\beta'})\delta_{\gamma,\gamma'}\Big(\frac{1}{8}(J(J+1)-k^2) \\ &+ \delta_{k,1}\frac{(-1)^{1-p}}{16}J(J+1)\frac{1+\gamma}{1-\gamma}\Big) \\ &+ V(r_{1\alpha},r_{2\beta},\theta_\gamma)\delta_{\alpha,\alpha'}\delta_{\beta,\beta'}\delta_{\gamma,\gamma'},\end{aligned} \tag{2.30}$$

where

$$K^{(i)}_{x,x'} = \sum_{t,t'} T_{xt} \langle t|\hat{h}^{(i)}|t'\rangle T_{x't'}, \tag{2.31}$$

$$L^{(i)}_{x,x',\gamma,\gamma'} = \sum_{l} T_{\gamma l} l(l+1) T_{\gamma' l} \sum_{t,t'} T_{xt} \langle t|\hat{g}^{(i)}|t'\rangle T_{x't'}, \tag{2.32}$$

$$M^{(i)}_{x,x'} = \sum_{t,t'} T_{xt} \langle t|\hat{g}^{(i)}|t'\rangle T_{x't'}, \tag{2.33}$$

from the quadrature approximation.

One of the attractive features of the DVR method is that, due to the quadrature approximation, the potential V is completely diagonal at every coordinate [138], and therefore requires no integration:

$$\begin{aligned} \sum_{m,n,l} \sum_{m',n',l'} & T_{\alpha m} T_{\beta n} T_{\gamma l} \langle m,n,l|V(r_1,r_2,\theta)|m',n',l'\rangle T_{\alpha' m'} T_{\beta' n'} T_{\gamma' l'} \\ & \simeq V(r_{1\alpha}, r_{2\beta}, \theta_\gamma) \delta_{\alpha,\alpha'} \delta_{\beta,\beta'} \delta_{\gamma,\gamma'}. \end{aligned} \tag{2.34}$$

Another advantage of the DVR3D method is that working within the DVR allows for the solution of the 3D Hamiltonian to be performed as a series of diagonalisations and truncations [138, 161], reducing the overall computational cost. In the current example using symmetrised Radau coordinates in a bisector embedding, the 2D Hamiltonian in the radial coordinates is constructed and diagonalised for a number of fixed angular coordinates. The lowest valued solutions to these intermediate calculations are then used in the construction of the full 3D Hamiltonian, a hierarchical scheme which is computationally efficient. Singularities at low angles may also be avoided by disregarding solutions to the 2D problem for small θ [161]. For each value of k the solutions of the 3D Coriolis-uncoupled Hamiltonian with the lowest energy are subsequently used to construct a fully Coriolis-coupled rovibrational Hamiltonian for a particular J, where the maximum number of these low energy solutions to be used is defined by the user. This final Hamiltonian therefore exhibits a diagonal k-block structure; using the solutions of the first step as a basis for this second step reduces the computation to that of determining the off-diagonal coupling terms only. The current work makes use of the ROTLEV3B module to perform this task, which uses a DVR for all three coordinates.

The final Coriolis-coupled rovibrational Hamiltonian has the form

$$\begin{aligned}\langle h,p,q|H|h',p',q'\rangle =&\delta_{h,h'}\delta_{k,k'}\varepsilon_h^{J,k}\\ &-(1+\delta_{k,0}+\delta_{k',0})^{-1/2}\delta_{k',k\pm1}\delta_{q',1-q}C^{\pm}_{J,k}\\ &\times\sum_{\alpha,\beta,\gamma,\gamma'}\psi^{J,k,h}_{\alpha,\beta,\gamma}\psi^{J,k',h'}_{\alpha,\beta,\gamma'}\left(M^{(1)}_{\alpha,\alpha,\beta,\beta}-M^{(2)}_{\alpha,\alpha,\beta,\beta}\right)J^{(2)}_{k\pm1,k,\gamma,\gamma'}\\ &-(1+\delta_{k,0}+\delta_{k',0})^{-1/2}\delta_{k',k\pm2}\delta_{q',q}C^{\pm}_{J,k\pm1}C^{\pm}_{J,k}\\ &\times\sum_{\alpha,\beta,\gamma,\gamma'}\psi^{J,k,h}_{\alpha,\beta,\gamma}\psi^{J,k',h'}_{\alpha,\beta,\gamma'}\left(\frac{1+\gamma}{1-\gamma}\right)\left(M^{(1)}_{\alpha,\alpha,\beta,\beta}+M^{(2)}_{\alpha,\alpha,\beta,\beta}\right)J^{(2)}_{k\pm2,k,\gamma,\gamma'}\\ &k=p,p+1,...,J,\qquad p=0,1,\end{aligned} \tag{2.35}$$

where q and p are defined as above. In Equation 2.35 the $|h\rangle$ and $\varepsilon_h^{J,k}$ are the eigenfunctions and associated eigenvalues, respectively, selected from the first DVR3DRJZ step for a particular J and k. The $C^{\pm}_{y,k}$ angular factors are given by

$$C^{\pm}_{y,k}=(y(y+1)-k(k\pm1))^{1/2}, \tag{2.36}$$

and the $J^{(i)}_{k',k,\gamma,\gamma'}$ are angular integrals which are given by

$$J^{(1)}_{k',k,\gamma,\gamma'}=\sum_{l,l'}T_{\gamma(k)l}\left[\delta_{l,l'}+\left(\frac{k+k'}{2}\right)\langle l,k|\frac{(1+cos\theta)}{sin\theta}|l',k'\rangle\right]T_{\gamma'(k')l'}, \tag{2.37}$$

and

$$J^{(2)}_{k',k,\gamma,\gamma'}=\sum_{l,l'}T^{\gamma(k)l}\left[\delta_{l,l'}+\langle l,k|\frac{(1+cos\theta)}{(1-cos\theta)}|l',k'\rangle\right]T^{\gamma'(k')l'}, \tag{2.38}$$

i.e. they are obtained by a DVR transformation from the corresponding FBR functions; the $\gamma(k)$ notation makes it explicitly clear that the γ DVR points are obtained from the Gaussian quadrature of the associated Legendre functions, which vary in k. Further details of this procedure may be found in Ref. [138].

2.5.4.1 The Final Rovibrational Wavefunctions

The resulting wavefunctions files generated from the ROTLEV3B calculations are structured according to k-blocks of the underlying basis functions [164], given as amplitudes at the DVR points. This is computationally advantageous for calculations of

the dipole transition strength where transitions are calculated according to $\Delta k = 0, \pm 1$ (see Section 2.8).

The wavefunctions are classified by the angular moment quantum number J, and the overall symmetry defined by the qp combinations of vibrational and rotational parities, which give rise to either A_1 or A_2 symmetry type under $\mathscr{C}_{2v}$.

2.6 The Nuclear Motion Problem: SO_3 & TROVE

The second half of the project described in this thesis focuses on the development of a line list for SO_3. Unlike for SO_2, where a specific code designed for the wavefunction construction of molecules of a particular structure (DVR3D for triatomic systems) is used, the rovibrational calculations for SO_3 have been performed with the program TROVE [131], which employs a variational approach in constructing rovibrational energies for general polyatomic molecules, using a numerically determined finite basis representation.

The calculations in TROVE can be performed in terms of any set of physically appropriate vibrational coordinates; in the current implementation linearised coordinates are used. The kinetic and potential energy operators are expanded numerically as a Taylor series in terms of these internal coordinates.

This section describes a general overview of the TROVE method, including a discussion of the important concept of symmetry involved with all calculations, particularly within the context of the SO_3 molecule. More detailed and general descriptions may be found in the published literature [131, 165].

2.6.1 The SO_3 Molecule

$^{32}S^{16}O_3$ is a stable, tetratomic molecule consisting entirely of spin-0 bosons. It's equilibrium configuration is that of a symmetric, trigonal planar structure, with the a sulphur atom located at the centre, and three Oxygen atoms bonded at 120 degrees apart within the plane [166]. It is this configuration that gives $^{32}S^{16}O_3$ the characteristics of the $\mathscr{D}_{3h}$(M) molecular symmetry group, which has six irreducible representations, defined in the $\mathscr{D}_{3h}$(M) character table [154]. Therefore, by determining the transformation properties of the wavefunctions computed for this molecule under the various operations of $\mathscr{D}_{3h}$(M), they can each be labelled by either A_1', A_2', E', A_1'', A_2'', or E''.

Functions representing the four fundamental modes ν_1 (symmetric stretch), ν_2 (symmetric bend), ν_3 (asymmetric stretch) and ν_4 (asymmetric bend) transform as A_1', A_2'', E' and E', respectively; the last two vibrational modes are therefore doubly-degenerate (see Fig. 2.5).

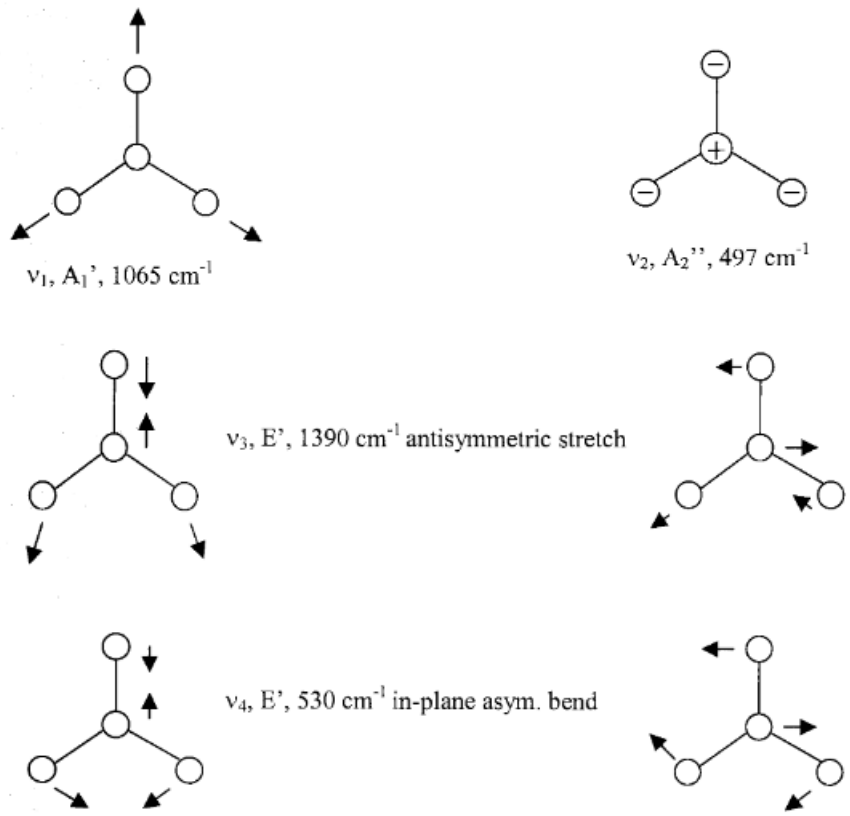

Figure 2.5: Fundamental vibrations of SO_3 (image taken from Ref. [166]).

As with SO_2, the Pauli Principle states that the total internal molecular wavefunction, Ψ_{Int}, must remain symmetric under the exchange of two identical Bosons. For $^{32}S^{16}O_3$ this is equivalent to a $\mathscr{D}_{3h}$(M) point group permutation which exchanges two of the Oxygen atoms. Analysis of the $\mathscr{D}_{3h}$(M) character table shows that only functions that transform as either A_1' or A_1'' obey this condition [154], with even and odd parity, respectively, and therefore only rovibrational wavefunctions described by these symmetry species need be constructed.

2.6.2 The Rovibrational Hamiltonian & Basis Set

The TROVE procedure concentrates on establishing a molecular Hamiltonian that is defined in terms of suitable coordinates within a molecule-fixed axis system. The kinetic and potential operators are then each expanded as a Taylor series in these coordinates

up to a certain order. The Hamiltonian, transformed into the molecule-fixed axis system can be written as [131, 165]

$$H = H_{vib} + \frac{1}{2}\sum_{\alpha,\beta} J_\alpha G_{\alpha,\beta} J_\beta + \frac{1}{2}\sum_{\alpha,n}(J_\alpha G_{\alpha,n}\frac{\partial}{\partial \xi_n} + \frac{\partial}{\partial \xi_n} G_{\alpha,n} J_\alpha), \tag{2.39}$$

where

$$H_{vib} = \frac{1}{2}\sum_{n,m}\frac{\partial}{\partial \xi_n} G_{n,m}\frac{\partial}{\partial \xi_m} + V + U \tag{2.40}$$

is the purely vibrational part of the Hamiltonian. The $J_\alpha(\alpha = x,y,z)$ operators are associated with the x,y and z -components of the total rotational angular momentum, and $\partial/\partial\xi_n$ are the vibrational momentum operators associated with the internal vibrational coordinate ξ_n. $G_{n,m}$, $G_{\alpha,\beta}$ and $G_{\alpha,n}$ are kinetic energy operators which depend on the internal coordinates , as do the potential and pseudopotential functions, V and U [131]. Each of these is expressed as an expansion in terms of the internal coordinates, and are implicitly dependent on the constituent atomic masses.

In this molecule-fixed Hamiltonian the energy due to the translational motion has been completely separated, leaving only terms describing the internal motion, due to vibration and rotation. The three components of the Hamiltonian in 2.39 essentially describe the pure vibrational motion, the pure rotational motion, and the coriolis interaction, respectively.

The aim is to express the kinetic and potential operators as Taylor series expansions in terms of the internal coordinates ξ_n, or functions thereof. TROVE allows the choice of any coordinates so long as it can represent the internal motions correctly. Linearised coordinates, or geometrically defined coordinates are examples of such coordinates that have been implemented in TROVE [131]. The expansion order of both of the kinetic and potential operators can be supplied by the user, and demonstrates the first example of the computational trade off between resource and quality of results; the truncation of each expansion is expected to introduce an error into the final results, and thus must be chosen to minimize the computational time while achieving the greatest accuracy (see Chapter 4).

2.6.2.1 The Vibrational Basis Functions

As previously stated, for TROVE, the rovibrational eigenfunctions of the molecular Hamiltonian are obtained in a finite basis representation (FBR) in terms of a specified basis set. Before constructing the full rovibrational basis set, it is necessary to consider the 'primitive' vibrational basis functions which are used in the first stage of computing the final rovibrational basis.

The aim is to define a purely vibrational basis composed of 1D primitive functions, each describing individual motions. TROVE employs a local mode approach [167]:

$$|v\rangle = \prod_n |v_n\rangle = \phi_{n_1}(\xi_{n_1})\phi_{n2}(\xi_{n_2})\phi_{n_3}(\xi_{n_3})...\phi_{n_{3N-6}}(\xi_{n_{3N-6}}). \tag{2.41}$$

Each ϕ_{n_i} is a function in the internal coordinate ξ_i, either defined as a standard analytical function (e.g. Morse oscillator or harmonic oscillator), or derived in a Numerov-Cooley numerical scheme [168, 169]. Each ξ_i describes a particular vibrational mode of the molecule; for a N-atomic non-linear molecule, $3N-6$ vibrational modes exist (for linear molecules, the number is $3N-5$). In the case of SO_3, six vibrational modes exist.

In the Numerov-Cooley scheme, the basis functions are obtained as eigensolutions to a 1-dimensional reduction of H_{vib} in Eq. 2.40 [131]:

$$H_{vib}^{1D} = \frac{\hbar^2}{2}\frac{\partial}{\partial \xi_n} G_{n,n}^{1D}(\xi_n)\frac{\partial}{\partial \xi_n} + V(\xi_n)^{1D} + U(\xi_n)^{1D} \tag{2.42}$$

For SO_3 a 6-dimensional vibrational basis set is created from various combinations of the 1-D functions. In theory, this basis set should be infinitely large in order to exactly describe the Hamiltonian matrix, however owing to the impracticalities associated with this, a truncation of the basis is required. TROVE incorporates a number of truncation schemes in order to define a basis cut off, whereby function combinations are kept or discarded based on their adherence to certain criteria, for example, only considering the products of functions whose individual 1-D function constituents have energies which sum to a value which lies below a certain threshold (see Chapter 4).

Like the truncation of the kinetic/potential energy expansions, an error is introduced to the final results due to the exclusion of various basis functions, and therefore

the truncation must be carefully fine-tuned in order to incorporate those which are necessary to ensure sufficient accuracy.

2.6.2.2 Symmetry Considerations in the Vibrational Basis Set

In order to facilitate further calculations a technique of 'symmetrisation' can be carried out whereby the purely vibrational Hamiltonian can be factorised into independent symmetry blocks [165, 170–172]. This can be thought of as a rotation of the vibrational basis, $|v\rangle$, into a set of symmetrised vibrational basis functions that can be classified under $\mathscr{D}_{3h}$(M) symmetry; in this approach, the Hamiltonian is diagonalised in terms of the new basis functions, $\phi_{n_i}^{\Gamma}$, where $\Gamma = A_1', A_2', E', A_1'', A_2''$, or E''.

This procedure greatly reduces the number of non-diagonal elements in the vibrational Hamiltonian matrix, however it does not eliminate them. As such, the resulting eigenfunctions will be linear combinations of the underlying basis functions. In order to characterise the final eigenfunctions by symmetry species, the symmetry label of the basis function with the largest contribution is used [172]. The symmetry labelling of the resulting vibrational eigenfunctions serves to simplify further calculations in conjunction with rotational basis functions where resulting combined rovibrational functions that are physically disallowed via the Pauli Principle may be ignored based on their symmetry properties, and computation of matrix elements may be reduced based on the vanishing integral rule.

2.6.2.3 The J = 0 Repesentation & the Final Rovibrational Basis Set

The resulting eigenfunctions computed as solutions to the 6-dimensional, symmetrized pure vibrational problem can themselves be used to form a compact basis set for the rotation-vibration solutions. This is the so-called $J = 0$ representation [165], $|\Psi_i^{J=0}\rangle$. Naturally, H_{vib} is diagonal in this basis

$$\langle \Psi_i^{J=0} | H_{vib} | \Psi_j^{J=0} \rangle = E_i^{J=0} \delta_{i,j}. \tag{2.43}$$

These eigenfunctions can be used in conjunction with rigid rotor functions to form the final rovibrational basis set [165, 171]:

$$|\Psi_{J,K,m,\tau_{rot}}^{i}\rangle = |\Psi_i^{J=0}\rangle \; |J, K, m, \tau_{rot}\rangle. \tag{2.44}$$

Here, $|J,K,m,\tau_{rot}\rangle$ is a symmetrised rigid rotor function, with τ_{rot} signifying rotational parity by $(-1)^{\tau_{rot}}$ [131]. Since H_{vib} is diagonal in the $|\Psi_j^{J=0}\rangle$ basis the matrix elements need not be computed in all further rovibrational calculations, which greatly simplifies the $J > 0$ problem. This is akin to the DVR technique previously discussed, by maintaining a diagonal vibrational basis set.

The matrix elements associated with the $G_{\alpha,\beta}$ and $G_{\alpha,n}$ operators in 2.39 can also be transformed into the $J = 0$ representation [165]; these can be precomputed and stored to disk for use in the rovibrational eigenfunction calculations which, along with the use of the $J = 0$ basis set, removes the need for further use of the larger primitive basis set.

Thus, the final rovibrational Hamiltonian evaluation is represented by [165]

$$\begin{aligned}\langle \Psi^i_{J,K,m,\tau_{rot}}|H|\Psi^{i'}_{J,K',m,\tau'_{rot}}\rangle &= E_i^{J=0}\delta_{i,i'}\delta_{K,K'} \\ &+ \frac{1}{2}\sum_{\tau_{rot},\tau'_{rot}}\sum_{\alpha,\beta}\langle J,K,m,\tau_{rot}|J_\alpha J_\beta|J,K',m,\tau'_{rot}\rangle \otimes \langle \Psi_i^{J=0}|G_{\alpha,\beta}|\Psi_{i'}^{J=0}\rangle \\ &+ \frac{1}{2}\sum_{\tau_{rot},\tau'_{rot}}\sum_{\alpha,n}\langle J,K,m,\tau_{rot}|J_\alpha|J,K',m,\tau'_{rot}\rangle \otimes \langle \Psi_i^{J=0}|G_{\alpha,n}\frac{\partial}{\partial\xi_n}+\frac{\partial}{\partial\xi_n}G_{\alpha,n}|\Psi_{i'}^{J=0}\rangle\end{aligned} \tag{2.45}$$

2.6.2.4 Symmetry Considerations in the Rovibrational Basis Set

The rotation of SO_3 can be approximated by using the rigid-rotor model for the case of a symmetric top [154]; in TROVE, the rovibrational basis is constructed as symmetrised combinations of such functions [131]. As with the vibrational basis functions, the symmetry of each rotor function may be determined by their behaviour under operations of the rotational subgroup of $\mathscr{D}_{3h}$(M).

Since the rovibrational wavefunction must have characteristics of either A_1' or A_1'' symmetry to obey the Pauli Principle, restrictions on the rotation-vibration combinations can be made. The product table for the $\mathscr{D}_{3h}$(M) point group shows the resulting characters of symmetry species combinations; the product of two or more irreducible representations results in a further representation which may be either itself irreducible, or a reducible species complex. Table 2.2 summarises the various species combinations which produce A_1' or A_1''.

The Hamiltonian for a specific J is constructed in the fashion of Equation 2.45 by

Table 2.2: Symmetry species combinations giving rise to A_1-type symmetry in the $\mathscr{D}_{3h}(M)$ symmetry group.

Γ_{vib}	Γ_{rot}	Γ_{Total}
A_1'	A_1'	A_1'
A_1'	A_1''	A_1''
A_2''	A_2''	A_1'
A_2''	A_2'	A_1''
E'	E'	$A_1' + A_2' + E'$
E'	E''	$A_1'' + A_1'' + E''$

combining the symmetrized rotor functions corresponding to the particular J with the symmetrized $J = 0$ vibrational basis. Thus, the procedure can be performed for each J individually.

The diagonalization of the Hamiltonian matrices for each J produces the wavefunctions and energies necessary for defining transitions and their corresponding strengths. These wavefunctions are labelled by symmetry species resulting from the combinations of the rotational and vibrational functions. Like in the case of the vibrational basis functions, the final wavefunctions are in fact linear combinations of the underlying basis functions, therefore their symmetry labelling (and, in fact, all quantum number labelling) result from analysis of which basis functions have the largest contribution.

For a given vibrational function, the combination of its symmetry species with that of the largest rotational contribution provides the total symmetry of the resulting wavefunction. Those that do not transform as A_1' or A_1'' are discarded.

2.6.2.5 The Final Rovibrational Wavefunctions

The rovibrational wavefunctions computed by TROVE are defined by a set of quantum numbers. Two rigorously conserved quantum numbers are the total angular momentum, J, and the overall symmetry, Γ_{Total}.

The remaining quantum numbers are approximate. For quantum numbers associated with rotation there is K, the projection of J on the the molecular z-axis, and is taken as the K associated with the largest contribution from the underlying rotational basis functions. The symmetry species associated with this rotational contribution, Γ_{rot} is also defined.

For vibrational quanta the local mode numbers are defined by the largest contribu-

tions from the underlying 1-dimensional, primitive vibrational basis functions. These local mode [173] quantum numbers can be mapped onto the commonly known normal mode quantum numbers of the Herzberg convention [174]. The vibrational symmetry associated with this contribution is given by Γ_{vib}.

2.7 Transition Intensity Simulations

The calculation of rovibrational wavefunctions represents the first half of a line list calculation. Already this provides sufficient information about the available energy levels accessible to the molecular system. The second half of the calculation involves determining the strengths of transitions between the energy levels.

2.7.1 The Dipole Moment

For a rovibrational transition to occur within a molecule an oscillating dipole moment must be induced by either rotation or vibration of the molecule, which can then interact with the oscillating electric field of incident electromagnetic radiation in order to enter another energetic state, where the difference in energy of the states is equal to the frequency of the interacting photon; this is known as an electric dipole transition. The dipole moment is a vector quantity describing the distribution of charge within a molecule. If a charge displacement exists across a molecule in its equilibrium configuration then it is said to have a permanent dipole moment, and vibrational distortions may also induce an instantaneous dipole moment due to asymmetry of the nuclear positions.

In the language of quantum mechanics the interaction between a charge distribution and an electric field can be represented by an 'interaction energy',

$$I = -\hat{\mu} \cdot \hat{E}, \tag{2.46}$$

where $\hat{E}$ is the external electric field, and $\hat{\mu}$ is the dipole moment, defined as

$$\hat{\mu} = \sum_i q_i \hat{r}_i, \tag{2.47}$$

where $\hat{r}_i$ is the position of charged particle i, given in arbitrary coordinates.

The strength of the interaction can be given by the expectation value of the inter-

action energy operator. Assuming that the electric field has constant magnitude over the duration of interaction, then the expectation value can be written as

$$\langle I \rangle = -\langle \psi_n | \hat{\mu} | \psi_m \rangle \cdot \hat{E}. \tag{2.48}$$

The case where $n \neq m$ is of interest, as these off-diagonal matrix elements give an indication as to the mixing of states ψ_n and ψ_m introduced by the perturbation I. In other words, if the matrix element is non zero, then the interaction induces transitions from ψ_m to ψ_n; the magnitude of the square of the integral is proportional to the probability, or strength of the transition.

The dipole moment operator $\hat{\mu}$ in this sense can be thought of as an analytical representation of a Dipole Moment Surface (DMS), which is an N-dimensional grid of dipole moment values as a function of the N degrees of freedom defining the molecular geometry. To this end, the DMS is structurally analogous to a PES, and is obtained in a similar fashion. Representation of a DMS in terms of appropriate internal coordinates (i.e. those used in the nuclear motion calculations) can help facilitate the calculation of the matrix elements described in Eq. 2.48.

2.7.2 Allowed & Forbidden Transitions

The discussion on symmetry in Subsection 2.4.3 provides a good precursor to the understanding of physically allowed molecular dipole transitions, particularly the concept of the Vanishing Integral Rule described in Eq. 2.15.

For example, by analysing the transformational properties of the dipole operator $\hat{\mu}$ under operations of the Molecular Symmetry group spanned by the eigenfunctions of the Molecular Hamiltonian, it is possible to determine which irreducible representation is generated by the operator. The vanishing integral rule states that the integral $\langle \psi_f | \hat{\mu} | \psi_i \rangle$ is only non zero if the representation generated by the integrand product, $\Gamma_f + \Gamma_\mu \otimes \Gamma_i$, contains the totally symmetric representation of the group.

It is easy to demonstrate that the dipole moment operator $\hat{\mu}$ is invariant under permutation operations of like particles, however it is antisymmetric upon inversion. As a consequence of this, $|\psi_i\rangle$ and $|\psi_f\rangle$ must have opposing parities [154]. Similar arguments in conjunction with the the rigid rotor functions for non-linear molecules show that electric dipole transitions allowed between states $|\psi_i^{J'}\rangle$ and $|\psi_f^{J}\rangle$ (which are

comprised of rigid rotor functions) are those where

$$\Delta J = J - J' = 0, \pm 1. \tag{2.49}$$

These are 'rigorous' selection rules defined for electric dipole transitions [154], which must also be obeyed in conjunction with

$$A_1 \leftrightarrow A_2 \tag{2.50}$$

for SO_2, and

$$A_1' \leftrightarrow A_1'' \tag{2.51}$$

for SO_3, in order to satisfy the parity change rule. These selection rules define the allowed transitions calculated by both DIPOLE3 and TROVE.

Other 'approximate' selection rules can also be enforced, however these are not based on well-defined quantum numbers. An example of this is for the K (z-axis projection of J) quantum number. Selection rules involving this number do not strictly forbid transitions, instead they usually define a strong transition. For example $\Delta K = 0$ is generally obeyed for strong transitions, however deformational effects often lead to K no longer being a good quantum number, and therefore $\Delta K > 0$ transitions can also occur, due to rotation-vibration coupling. Such behaviour is the source of the purely rotational transitions of SO_3, which at equilibrium has no permanent dipole moment. Chapter 5 discusses this effect at length.

2.7.3 Energy State Population & Temperature Dependence of Transition Intensities

The transition dipole moment gives a quantum mechanically 'intrinsic' measure of the strength of a molecular transition. However, the physical manifestation in reality of transitions allowed via the vanishing integral rule depends on the population of the states involved.

2.7.3.1 Degeneracy of the Energy Levels

In the absence of an external magnetic field there exists a systematic degeneracy of the rotational levels. The quantized angular momentum is aligned along a total of $2J + 1$ spacial orientations, where J is the angular momentum quantum number associated with the state in question. The population of an energy state will depend on this degeneracy factor, and therefore calculations must be weighted accordingly.

Another source of degeneracy may arise from the nuclear spin statistics [154]. This is best understood by a separation of the total wavefunction into electronic, rovibrational and nuclear spin components:

$$\Psi_{Total} = \Psi_{Electronic}\Psi_{Rovib}\Psi_{NucSpin}. \tag{2.52}$$

The total wavefunction Ψ_{Total} must obey the Pauli Principle in terms of Fermi-Dirac or, in the case of SO_2 and SO_3, Bose-Einstein Statistics; for the case of SO_2 and SO_3 the wavefunction must remain symmetric under exchange of identical Bosons (Oxygen). This puts restrictions on the physically allowed combinations on the right-hand side of Eq. 2.52 as their product must generate a symmetric irreducible representation.

For the work described in this thesis, only rovibrational motion within the electronic ground state of both SO_2 and SO_3 is considered, both of which are fully symmetric closed shell systems, i.e., $\Gamma_{Electronic}$ transforms as the fully symmetric species of the appropriate point groups for each molecule [154]. Therefore the product $\Psi_{Rovib}\Psi_{NucSpin}$ must ensure that Ψ_{Total} is fully symmetric under the exchange of identical Bosons.

The Bosons in question, the Oxygen nuclei, have zero spin and therefore only a single, symmetric nuclear spin configuration is possible for all combinations of $\Psi_{Rovib}\Psi_{NucSpin}$, meaning that Ψ_{Rovib} must also be a symmetric function. The degeneracy factor arising from the nuclear spin statistics is therefore simply 1 for all Ψ_{Total} as there are no other nuclear spin configurations that occur.

This represents a rather simple case whereas in other cases such as those involving two Hydrogen nuclei, due to the Fermionic nature of the particles, Ψ_{Total} would need to be antisymmetric upon their exchange, and since they also posses $\frac{1}{2}$-spin a total of three symmetric and one antisymmetric nuclear spin states become possible, which results in

a 3:1 ratio of 'ortho' and 'para' states when combining with Ψ_{Rovib}.

2.7.3.2 The Effect of Temperature

The characteristic features of a molecule's spectrum will vary with the absolute temperature at which it is observed. The individual transitions within a spectrum between pairs of states will have an intensity that is proportional to the probability of the occupation of the initial state, which is given by a Boltzmann factor:

$$P(E_i) \propto e^{-E_i/k_bT}. \tag{2.53}$$

Here, E_i is the energy of the state i, k_b is the Boltzmann constant, and T is the absolute temperature. Clearly, the population of the state increases with the temperature of the sample, hence the transitions originating from the state in question increase in number, and therefore so do their intensities. Each probability must be weighted by the degeneracy factor of the state E_i to obtain an accurate estimate of the level's population. To ensure normalization for the probability $P(E_i)$ a weighting factor may be introduced as a probability summation over all possible energy states:

$$Q = \sum_i g_i \, e^{-E_i/k_bT}. \tag{2.54}$$

This value is known as the partition function, and is an important thermodynamic quantity. The g_i represents the total degeneracy of the state E_i, which for a rovibrational level is equal to $g_{ns}(2J_i+1)$, where g_{ns} is the nuclear spin degeneracy. The accurate calculation of the partition function for a given system is vital for determining the occupation probabilities of states.

In reality, summing over all possible states of the system is not practical as the energies may not necessarily all be calculated, particularly higher ones. Provided that the partition function is adequately converged up to a given energy threshold, then the summation may be truncated. However as the temperature increases in Eq. 2.54, higher energies are required for convergence of Q; they 'contribute more' to the partition function, which is physically equivalent to stating that these higher energies become more occupied at higher temperatures.

2.7.3.3 Intensity Simulations

The intensity of an absorption line, assuming thermal equilibrium at the absolute temperature T, is related to the line strength $S(f \leftarrow i)$ [165]:

$$I(f \leftarrow i) = \frac{8\pi^3 N_A \nu_{if}}{(4\pi\varepsilon_0)3hc} \frac{e^{-E_i/k_b T}}{Q} \times [1 - e^{-hc\nu_{if}/k_b T}] S(f \leftarrow i). \tag{2.55}$$

Here, ν_{if} is the wavenumber of the line, between energy states E_i and E_f, N_A is Avogadro's constant, h is Planck's constant, c is the speed of light in a vacuum, k_b is Boltzmann's constant, and ε_0 is the permittivity of free space. To ensure accurate intensity results for any absorption line at the absolute temperature, T, it is vital that the partition function, Q, is adequately converged, as discussed above.

The temperature-independent part of the intensity calculation is encapsulated by $S(f \leftarrow i)$, the transition strength. The general definition of this quantity is of a matrix element of the dipole operator, the derivation of which depends on the structure of the rovibrational wavefunctions. In Ref. [165], in a accordance with the TROVE method, it is given by

$$S(f \leftarrow i) = g_{ns} \sum_{m_f, m_i} \sum_{A=X,Y,Z} |\langle \Psi^f_{rv} | \hat{\mu}_A | \Psi^i_{rv} \rangle|^2, \tag{2.56}$$

where g_{ns} is the nuclear spin statistical weight factor [154], μ_A is the electronically averaged component of the molecular dipole moment along the space-fixed axis $A = X, Y, Z$; μ_A is expressed in terms of the dipole moment components along the molecule fixed axes. The quantum numbers m_i and m_f are the projections of J onto the Z axis for the initial and final states, respectively. The $|\Psi^i_{rv}\rangle$ and $|\Psi^f_{rv}\rangle$ are the rovibrational states associated with the initial and final states. The line strength may be computed in a number of different ways [138] (see Subsection 2.8.3).

When producing a line list it is useful to list only the temperature-independent form of the transition intensity, as any temperature-dependent result may be extrapolated from this, for example via Eq. 2.55.

2.7.4 The Final Line List

The structural form of a final line list requires both information for a transition (i.e. its strength), and for the states between which the transition occurs. To this end, the ExoMol line list format consists of two files [175].

The first list is a states file listing all the rovibrational energies calculated for the molecular system. Each energy state is labelled according to a set of quantum numbers produced from the calculations. 'Rigorous' quantum numbers include the energy value, the rotational quantum number J, and the state degeneracy. Further labels may be introduced to ensure the assignment of states is less ambiguous; these include parity, further rotational quantum numbers, e.g. K, vibrational quantum numbers (detailing the level of excitement in the vibrational modes), and symmetry labels for the total wavefunction, and its rotational and vibrational constituents. The extent of the quantum number labelling is dependent on the method of the rovibrational wavefunction calculation. For example, with TROVE it is possible to label states according vibrational local (and normal) modes according to the vibrational contribution to the state, as local mode functions are an inherent feature of the TROVE calculation. Conversely, DVR3D does not produce any vibrational labels. Nevertheless, each level in the file may be uniquely identified via a running index, which is made use of in the second line list file.

The second file, the transition file, is more straightforward. It contains a list of transitions labelled by the indices of the initial and final states in accordance with the states file, and the strength of the transition between these two states, resulting in a total of three columns. The strength of the transition is given in the form of the Einstein A-coefficient, which is temperature-independent and related to the quantity $S(f \leftarrow i)$ in Eq. 2.56. This ensures that any temperature-dependent simulation may be extrapolated from the line list for a whole range of T.

The work in this thesis consists of the calculation of various line lists which are suitable for different temperatures. Since the line lists do not explicitly specify a temperature, it is important to explain what is meant by a 'room-temperature line list' and a 'hot line list'. A line list is suitable for modelling spectra up to a temperature T if all the energies that are significantly occupied at this temperature have been calculated. This can be checked with the partition function, Q, in Eq. 2.54; if this quantity is converged for the absolute temperature T then the energy level coverage from all the E_i can be assumed to be sufficient. A line list containing transitions from all these E_i is therefore suitable for describing spectra at the given temperature.

2.8 Improvements to the DVR3D Suite for the Computation of the SO_2 Line List

Calculating a molecular line list can be a formidable task and depends on a multitude of factors. The ultimate aim is to compute wavefunctions and energies suitable for determining strengths of all transitions which are significant for a given spectral range at the desired temperature. This requires not only a suitable number of rovibrational energies, but the additional requirement that they are adequately converged using a basis set of an appropriate size. The optimisation and truncation of the initial basis set has an effect on all subsequent calculations; the computation time, memory requirements and hard disk storage space are all dependent on the size of the initial basis. Estimating the amount of these resources required for the calculation of the complete linelist is not straightforward.

Although constructed for the purpose of triatomic rovibrational calculations, the DVR3D suite was not readily optimised for the the SO_2 problem at the start of the project due to the requirements of basis set convergence for this molecule. Initial calculations quickly demonstrated that various adaptations to the code were required in order to deal with the large basis set, and to enable computations for wavefunctions and transition moments for SO_2. The changes made to each module are now described.

2.8.1 DVR3DRJZ

The changes made to the first module of the program suite focus on various aspects of the gaussian quadrature calculations.

Changes were required for a routine '*ASLEG*' (associated Legendre) that computes the associated Legendre functions used in building the angular basis set. The routine by Press *et al* [176] encounters issues with surpassing limits of double precision in the IEEE Standard for Floating-Point Arithmetic; namely, for high rotational quantum number J (and subsequently k), the routine attempts to compute numbers that are beyond the the limit for the 64 bit number format. In the Fortran standard, the function $HUGE(x)$ returns the largest number that is not an infinity of the type x. For the 64-bit Intel compiler this number is roughly equal to 2×10^{308} in double precision.

For values of $k > 90$ this proved to be an issue in the original implementation of

the routine by Press *et al* and would lead to computing infinities over various loops. To prevent this it was necessary to implement some conditional statements to reduce the size of these numbers to work with the original machinery of the routine by using a constant scaling factor, which is chosen to ensure that subsequent growth of these numbers over DO loops is restricted to be within the range of the double precision format. Calculations are subsequently performed using these numbers and final results are then multiplied once more by the original scaling factor to achieve the desired values with a negligible change in accuracy. This modification routine was also identically implemented into the ROTLEV3B and DIPOLE3 modules of the DVR3D suite where the computation of associated Legendre polynomials is also required, and therefore will not be addressed in the following subsections.

The second change change deals with a subroutine to calculate the zeros of the highest-order associated Legendre polynomial and corresponding weights to be used in the Gauss-associated Legendre quadrature scheme for the radial functions [177]. The subroutine computes the quadrature points using a brute force search for the zeros of the N^{th} order polynomial (where N is an argument passed into the subroutine) by stepping through the polynomial and looking for a change in sign.

Beginning at $x = 1$, the routine steps through the polynomial at equal intervals predefined in the code. In the original code design the stepping interval was taken to be 10^{-4}, and the value of the polynomial is calculated at each value of x stepping backward from $x = 1$. When a change of sign is found then the zero of this N^{th} degree polynomial is taken to be the value of x at that point. This is done until all the zeros are found (in fact, only the zeros for the region $0 < x \leq 1$ are required which constitute half of the total amount, the remainder of which can be easily found as the polynomials are symmetric about 0).

Initial convergence tests of the basis set for this work (see Chapter 3) could not be achieved using this routine, and would fail before any convergence was observed. In order to be able to calculate the necessary quadrature points, the code was adapted first by decreasing the size of the stepping interval to 10^{-6}, and then by defining the zero to be the mid-point between the current value of x and the previous value. This change has proven adequate to suit the requirements.

2.8.2 ROTLEV3B

This module remains largely unchanged other than for the '*ASLEG*' routine previously described. This is required in ROTLEV3B in order to calculate associated Legendre functions in accordance with Equations 2.37 and 2.38. It was also found that, for high J, vibrational basis functions generated by the DVR3DRJZ run are not required above certain values of k to construct the Hamiltonian, whose dimension is defined by an input parameter (see Chapter 3). Given that these are associated with high energies, they are therefore discarded, and the '*ASLEG*' routine is not called for these values of k in order to optimise time usage. The main difference in the ROTLEV3B module is in how the final rovibrational wavefunctions are saved, which has proven essential for ensuring that they are stored completely, and in order to perform the computation of transition moments later on using DIPOLE3. Although the modification only requires a simple change to some Fortan '*open*' statements, it is worth discussing the reasoning and purpose.

The wavefunction files written by ROTLEV3B are unformatted which reduces storage requirements and I/O usage in further processing. The layout of the wavefunctions also optimises I/O by being structured into blocks of the k quantum number (projection of J onto the body-fixed z-axis). Each record within the file corresponds to a k-block containing data fields for each energy level. This structure is used advantageously in DIPOLE3 by reading through the k-block records sequentially to compute transition moments.

The standard way in which these wavefunctions are saved are in the Fortran variable-length record type. Over the course of the work it became apparent that, for the SO_2 problem in the chosen basis set, this type of sequential file organisation failed as the records being written were surpassing the record length limit imposed by the operating system. To resolve this issue it is possible to write records in a 'segmented' structure. The segmented record type is unique to Intel Fortran. A record written in this style consists of multiple variable-length, unformatted records (segments) which together combine into the one single logical record. The "*segmented*" flag can be used to write records where it is not desired or not possible to write a single-variable length record. For this work the segmented record type facility was used to store all wave-

function files, and assures the completeness of each record and, particularly important for this work, that they are read in fully by the DIPOLE3 module.

2.8.3 DIPOLE3

The main issues that lead to the modifications of the DIPOLE3 module were related to I/O performance. The DIPOLE3 procedure is outlined in Ref. [138] and in greater detail in Ref. [178]. DIPOLE3 is already successfully implemented using OpenMP for parallelisation across multiple CPUs, which speeds up performance of the BLAS routine '*DGER*' for the necessary transformations which were the original source of maximum CPU usage [138]. However, for the current work on SO_2 the module must read in extremely large wavefunction files computed and stored by ROTLEV3B (peaking at around 0.5 TB), and the existing method for reading records proved to be a bottleneck in the computation of transition moments.

DIPOLE3 performs the following calculation for the linestrength [138]:

$$\begin{aligned} S(f \leftarrow i) =& \frac{1}{4}[(2J'+1)(2J''+1)][(-1)^{J''+J'+1}+(-1)^{p'+p''}]^2 \\ & \times \left[\sum_{\nu=-1}^{+1}\sum_{k=p''}^{J''} a(k+\nu,k)(-1)^k \begin{pmatrix} J' & 1 & J'' \\ -\nu-k & \nu & k \end{pmatrix} \sum_{\alpha\beta\gamma} c^{J'p'x'}_{k'\alpha\beta\gamma} c^{J''p''x''}_{k''\alpha\beta\gamma} \mu^m_\nu(\alpha\beta\gamma)\right]^2 \end{aligned} \tag{2.57}$$

between initial and final states x'' and x', with J'', k'', p'' and J', k', p' quantum numbers, respectively. Here, the c^{Jpx} are the values of the wavefunction of the xth state at the DVR grid points specified by α, β and γ, and $\mu^m_\nu(\alpha\beta\gamma)$ is the value of the molecule-fixed dipole moment, specified on the same grid. The first sum runs over ν, which is equal to zero for a dipole moment aligned along the z-axis, and ± 1 when there is a dipole component across the x-axis in either direction. The second sum runs over the k quantum number; the c^{Jpx} are k-dependent as the ROTLEV3B wavefunctions are constructed from the k-dependent DVR3DRJZ basis set.

The phase factors in Equation 2.57 give the already known rigorous section rules:

$$\begin{aligned} \Delta J = 0, \qquad & p' \neq p'', \\ \Delta J = 1, \qquad & p' = p''. \end{aligned} \tag{2.58}$$

In addition to these, as a consequence of the Wigner 3-j symbols the further rules are defined as $\Delta k = 0, \pm 1$, depending on the value of v. Thus, DIPOLE3 performs all linestrength calculations for initial and final states satisfying these rules. The process involves navigating through wavefunction files and extracting the necessary information to perform the above calculation.

The wavefunctions produced by ROTLEV3B are stored in a k-block structure in terms of the original DVR3DRJZ basis, which are given as amplitudes at the DVR points. However the angular DVR points are k-dependent, and therefore DIPOLE3 first transforms all the ROTLEV3B wavefunctions onto a common DVR grid [138, 178], where the transformation is defined by the Gauss-Legendre quadrature of a zero-order polynomial (k = 0), defined by the number of angular points. The resulting wavefunctions are the c^{Jpx} in Equation 2.57. The advantage of the wavefunction storage by the k-block structure is due to the $\Delta k = 0, \pm 1$ selection rules previously mentioned. For a given transition there are two wavefunction files to be accessed, one containing the bra (initial state) wavefunctions and the other the ket (final state) wavefunctions. First, the coefficient data records of each file are read through simultaneously and sequentially; for each k-block the k-dependent wavefunctions are read and used subsequently in the main section of the code that deals with computing the transition moments, i.e. Equation 2.57 (this is the section where '*DGER*' is called). These $k' \leftarrow k''$ transition moment calculations for $\Delta k = 0$ are performed for all $k \leq J$.

This is followed by a '*rewind*' of the wavefunction files to begin from the start of the k-block records in order to calculate transition moments for $\Delta k = \pm 1$. This involves moving through the ket wavefunction file one k-block at a time for each k', as before, but by switching back and forth through the bra wavefunction files to access the $k'' = k' - 1$ and $k'' = k' + 1$ records by using a '*rewind*' statement for each cycle. This has proved to be highly inefficient for reading through the SO_2 wavefunctions computed with ROTLEV3B. These wavefunction files are so large that I/O performance with DIPOLE3 using the described method is significantly worse than for smaller problems, and it is not possible to compute the necessary transition strengths in a reasonable amount of time.

To deal with this problem available memory resources on computing nodes have been employed; the current work makes use of the High Performance Computing Ser-

vice at Cambridge, specifically the Darwin cluster, where each compute node provides up to 64 Gb of RAM which can be utilised for each computation. The solution therefore lies in taking advantage of this resource by storing the bra wavefunction records in the memory instead of navigating to them using the I/O performance-degrading '*rewind*' function. Records from the bra file can be stored in memory arrays and cycled through sequentially. In this way, the bra wavefunction file is read through only once in an identical way to the ket file, in order of increasing k. The required RAM to store a single k-block record from a wavefunction file for a given J is roughly equal to the size of the file divided by J (since there are J amount of k-blocks). Using this method there are four records stored in the memory at any given time (one from the ket file, three from the bra file), which would mean that roughly $(4\times$ size of J' wavefunction file$)/J'$ Gb of RAM are required for the calculation. As an example from the current work on SO_2, wavefunction file size peaks at around 0.5 Tb for $J = 60$, and therefore a DIPOLE3 calculation with this new implementation would require around 33 Gb of RAM, which fits comfortably within the 64 Gb memory resource for each computing node.

This has improved I/O performance immensely for the current work on SO_2, allowing for dipole transition moment calculations to be carried out fully within the 36 hour wall clock limit on the HPCS Darwin system.

Chapter 3

A Hot Line List for $^{32}S^{16}O_2$

In order to compute a line list for $^{32}S^{16}O_2$ using the DVR3D program suite [138] discussed in Chapter 2 three things are required: a suitable potential energy surface (PES), dipole moment surface (DMS), and an adequately converged basis set. The PES and the DMS used in this work are those of Huang *et al.* used in their production of a room-temperature line list [35].

In this chapter, the production of the hot line list is presented. The various convergence tests for the basis set are discussed, along with an outline of the computational approach of the ro-vibrational wavefunctions and subsequent transition moments. Problems encountered within the different aspects of the computation are also touched upon. The finalised hot line list is analysed and compared to some available experimental and theoretical data within the literature, in order to assess the reproducibility of the DVR3D calculations. Simulations of hot spectra are also compared to recently obtained experimental intensity cross sections.

3.1 The AMES SO_2 Line List

Huang *et al.* reported a highly accurate PES, DMS and IR line list for $^{32}S^{16}O_2$ up to 8000 cm^{-1} [35], computed at 296 K (denoted Ames-296K). The methods follow closely those reported in their works on H_2O [179], NH_3 [180, 181], and CO_2 [182]. Both the PES and DMS were computed at the CCSD(T) [115] level of theory.

The energy level coverage includes eigenstates up to 13 168 cm^{-1}, for all $J \leq 80$, computed using the program VTET [179, 183], using an *ab initio* CCSD(T)/cc-pVQZ-DK [184] computed PES later refined based on experimental data provided by HITRAN [75]. In total, 30 129 energy states were used in the refinement. The root

mean square (RMS) error for all $J \leq 80$ energy levels was 0.013 cm^{-1}. Transition intensities were computed based on an *ab initio* CCSD(T)/aug-cc-pV(Q+d)Z [184,185] computed DMS.

The aim of the current work is to extend upon the Ames-296K line list and produce a hot line list based upon the same PES and DMS, by computing energies up to 15 000 $cm^{-1}$and beyond the $J \leq 80$ limit, for the $0 < \nu \leq 8000$ cm^{-1} region. The following section discusses the implementation of the PES and DMS into the DVR3D program suite [138], as well as an overview of the computational process.

3.2 Calculation using DVR3D

The procedure for producing a line list with the DVR3D program suite consists of four major steps:

DVR3DRJZ This first module solves the vibrational problem, either for the purely vibrational case ($J = 0$) or for an initial rotationally-uncoupled vibrational basis to be used later in ROTLEV3B ($J > 0$). For each J, two runs of this module must be performed to obtain two set of wavefunctions, labelled by by even vibrational symmetry ($q = 0$) and odd vibrational symmetry ($q = 1$), respectively. The DVR3DRJZ module requires a number of inputs to produce these vibrational wavefunctions. Firstly, an accurate potential energy surface must be supplied to the code. Secondly, a set of parameters to define the structure of the DVR basis set, specified as the amount of DVR grid points in the radial and angular coordinates, as well as the Morse oscillator parameters discussed in Chapter 2. The optimisation of all these input parameters will be discussed in the next section.

ROTLEV3B In this module the vibrational wavefunctions for ($J > 0$) are taken from DVR3DRJZ as a compact basis in order to calculate the fully ro-vibrational wavefunctions for a certain vale of J. For each value of q in DVR3DRJZ there are two rotational parities that are computationally allowed, $p = 0$ (even) and $p = 1$ (odd), giving rise to a total of four possible combinations of symmetry labelling for the final ro-vibrational wavefunctions. However it is only necessary to compute the wavefunctions of two of the four possible combinations for a given J (which ones depend on whether J is even of odd), as certain combinations are physically

disallowed due to symmetry arguments. ROTLEV3B also requires the input of some optimisation parameters, which will also be discussed alongside those of the DVR3DRJZ module.

DIPOLE3 Two outputs from ROTLEV3B are taken as inputs to this module, which computes the dipole transition strengths (Einstein A coefficients) of all possible transitions between the bra and ket wavefunction files, provided that they obey rigorous, symmetry-based selection rules. The DIPOLE3 inputs allow the user to specify the number of upper and lower energies to be considered in a run, which can be important in reducing the computational time for a single calculation and enabling the process to be separated and run in parallel, with each job calculating for different upper and lower energy ranges.

SPECTRA Finally, this module takes the outputs from DIPOLE3 to calculate temperature-dependent line strengths of any transition computed in the previous module. The temperature of interest is specified as an input parameter, amongst others such as energy and J cut-offs to produce a temperature-dependent spectrum. The partition function value can either be calculated using this module, or specified as an input in order to produce a spectrum. Technically it is possible to produce a 'final' line list using only data computed in the DIPOLE3 module, as this contains all the information of energies, transitions and Einstein coefficients, which can be later used in subsequent programs built to interpret the DIPOLE3 data as an input.

3.2.1 Convergence Testing

The basis functions used in the DVR3D program suite are discussed in Chapter 2; for the angular basis functions Legendre polynomials are employed, and for the radial basis functions employ Laguerre polynomials in the form of Morse oscillator-like functions. The truncation of the combined basis set is controlled by a number of parameters which are supplied to the user input of the various modules of the DVR3D code. The procedure involves an empirical trial and error approach to obtain the optimal values for these parameters, to ensure the initial basis set is as accurate as possible, while minimising computational cost.

The majority of these control parameters are entered into the first module DVR3DRJZ, and define the DVR grid points upon which the rovibrational wavefunctions are constructed. Two parameters, NPNT2 and NALF define the number of DVR grid points in the radial coordinates and angular coordinate, respectively, from their associated quadrature schemes. The NPNT2 parameter actually defines the number of points in both radial coordinates due to the symmetry of the SO_2 system. The Radau grid must be optimised by a trial scheme to determine the best values of these parameters which obtains adequate convergence.

3.2.1.1 Parameters in DVR3DRJZ

The convergence testing within DVR3DRJZ calculations is essentially a minimisation scheme for the vibrational term values. Therefore, the Morse oscillator-like functions that represent the basis functions isomorphic to the DVR require optimising. This involves choice of the values of r_e, D_e, and ω_e (see Chapter 2), as well as the atomic masses. These were chosen as $M_{Sulphur}$ = 31.963294, and M_{Oxygen} = 15.990526, in atomic mass units.

In general, it is found that results are insensitive to the variation of D_e [138], and it is only r_e and ω_e which require variation. However these are strongly coupled and cannot be independently optimised, therefore convergence testing in these parameters involves analysing results across a grid of combinations.

Ultimately, the values of $r_e = 3.0$ and $\omega_e = 0.005$ yielded an adequate minimisation in energies for $J = 0$ tests. It should be noted that these values are not physical in that they do not necessarily represent those of the 'true' vibrating bond in question. In fact, the optimal value of r_e is generally found to be larger than the value obtained from the minimum of the Morse potential [138]. They are instead treated as empirical parameters to be determined.

The next parameters to be determined are those that determine the number of DVR points required in both the radial and angular coordinates $NPNT2$ and $NALF$, respectively, i.e. the number of points required for their associated quadrature schemes. The choice of these values actually has an effect on the optimisation of the Morse parameters, and therefore they are set to high, computationally expensive values during the previous phase, with a view to reducing them. The values $NPNT2 = 30$ and $NALF =$

130 are well suited for adequate convergence within reasonable computation time, and are therefore chosen for the remaining calculations.

Finally, the dimension of the Hamiltonian matrix, *MAX*3*D*, which also affects convergence must be specified. Like with *NPNT*2 and *NALF* this may be initially set to a high value in order to determine the Morse oscillator parameters and then reduced to speed up computation time, whilst retaining the accuracy of the final results. *MAX*3*D* is set to 3052 for all calculations using DVR3DRJZ. The number of eigenvalues/eigenvectors retained from this calculation to be used in ROTLEV3B is chosen as *NEVAL* = 1000. The optimisation of this parameter is actually performed in conjunction with ROTLEV3B calculations, as the number of DVR3DRJZ functions used in further calculations affects the results of the final rovibrational eigenvalues.

3.2.1.2 Parameters in ROTLEV3B

Once the DVR3DRJZ parameters are found to be suitably optimised for basis set convergence, it is necessary to do so for the ROTLEV3B module, and this includes finding the best value for the size of the Hamiltonian. This value, *IBASS* is dependent on J via the relation $IBASS = n(J+1)$; the Hamiltonian is built from the selection of the *IBASS* lowest energies (eigenfunctions) for each k computed by DVR3DRJZ. The testing process in this case consists of empirically determining a value of n such that the required ro-vibrational eigenvalues are converged after diagonalisation of the Hamiltonian. For this work convergence tests for *IBASS* are performed for J = 5, J = 10 and J = 60.

The determination of n is important since, like NPNT2 and NALF the computation time increases as the value of n increases, therefore it is important to find the minimum value for which adequate convergence is observed. The higher energy values are much more sensitive to this value, and therefore the optimum value of n is found by observing the convergence behaviour in this region. For SO_2 it is found that a value of n = 725 shows very good convergence of energies up to 15 000cm^{-1}. Table 3.1 demonstrates the level of convergence by comparing results computed with values of n above and below 725. The input parameters for DVR3DRJZ and ROTLEV3B calculations are summarised in Table 3.2.

Table 3.1: Convergence of rovibrational energies. Tests are performed for various values of J; here, for $J = 60$, the sum of energies below 10 000, 11 000, 12 000, 13 000, 14 000, and 15 000 cm^{-1} are listed for various values of n. Convergence is observed in all cases shown for energies below 10 000cm^{-1}.

J	E_{max}/ cm^{-1}	$n = 600$	$n = 650$	$n = 700$	$n = 750$
60	10 000	33060146.8373	33060146.8373	33060146.8373	33060146.8372
	11 000	50025398.5194	50025398.5188	50025398.5185	50025398.5184
	12 000	73813352.9930	73813352.9852	73813352.9818	73813352.9809
	13 000	104959184.7657	104959184.6718	104959184.6316	104959184.6202
	14 000	145689245.2379	145689244.1034	145689243.6767	145689243.5457
	15 000	197570330.8183	197570317.0376	197570312.5276	197570311.0017

Table 3.2: Input parameters for DVR3DRJZ and ROTLEV3B modules.

DVR3DRJZ		
Parameter	Value	Description
NPNT2	30	No. of DVR points in the radial coordinates (Gauss-Laguerre)
NALF	130	No. of DVR points in the angular coordinate (Gauss-Legendre)
NEVAL	1000	No. of eigenvalues/eigenvectors required
MAX3D	3052	Maximum dimension of the final Hamiltonian
XMASS (S)	31.963294	Mass of Sulphur atom in amu
XMASS (O)	15.990526	Mass of Oxygen atom in amu
r_e	3.0	Morse parameter (radial basis funciton)*
D_e	0.4	Morse parameter (radial basis funciton)
ω_e	0.005	Morse parameter (radial basis funciton)
ROTLEV3B		
Parameter	Value	Description
NVIB	1000	No. of vib. levels read in from DVR3DRJZ for each k
n	725/500§	Defines $IBASS = n(J+1)$ (see text)

*See Chapter 2.

$^\S n = 725$ for $J < 124$, and $n = 500$ for $J \geq 124$ (see text).

3.2.2 The Potential Energy and Dipole Moment Surfaces

As has been previously stated, in order to calculate accurate line positions and transition strengths it is necessary to have a good description of the potential energies and dipole moments, respectively, as a function of molecular geometry. The electronic structure methods involved in obtaining these surfaces are numerous and with varying degrees of accuracy, some examples of which are described in Chapter 2. Producing a functional form of both the potential energy surface (PES) and the dipole moment surface (DMS) that are complementary to solving the nuclear motion and dipole transition problems is necessary for their implementation into a variational rovibrational code,

such as DVR3D.

In this subsection the potential energy surface and the dipole moment surface computed by by Huang *et al.* [35] will be discussed, including overviews of their calculation, and their correct implementation into the DVR3D code.

3.2.2.1 The AMES-1 PES

The Potential Energy Surface initially used in this work is that presented by Huang *et al.* [35], which is used in the computation of their room-temperature "Ames-296K" line list. This *ab initio* potential is constructed using a coupled-cluster approach for singles and doubles, with perturbative triples (CCSD(T) [115]) for the electronic structure calculations across a total of 489 geometries. The correlation consistent basis set employed was cc-pVQZ-DK [184], chosen after comparing the deviations from 125 experimental energies with theoretical energies produced from 26 different PESs associated with various other basis sets. The average deviation of the compared energies was 5.34cm^{-1} for the cc-pVQZ-DK basis set. The points are fitted to a global function consisting of 219 coefficients.

The PES (which will be referred to as AMES-1) is presented as a subroutine with an associated coefficient file. This subroutine must be implemented into the DVR3DRJZ code; it returns the value of the potential, V, corresponding to the configuration of Radau coordinates R_1, R_2, and $XCOS$ supplied as arguments to the subroutine, where the bond lengths are in Bohr and $XCOS = cos\theta$.

After convergence testing, as described in the previous section, comparisons for the $J = 0$ problem were carried out between the DVR3RJZ computed vibrational energies and those published by Huang *et al.* [35] in order to ensure that the PES was implemented correctly. These comparisons appeared to show some discrepancies.

Huang *et al.* published values for their calculated vibrational bands up to around 5000 cm^{-1} using their variational rovibrational program VTET [179, 183], and labelled the band origins with normal mode quantum numbers. These were used as the basis of comparison for the corresponding energies computed with DVR3DRJZ. The major source of discrepancies appeared in bands associated with the ν_3 vibrational quantum number, which describes the asymmetric stretch normal mode.

Based on this disagreement the initial reaction was to re-assess the DVR3DRJZ

basis used for calculating the vibrational energies, the atomic masses assumed, and to ensure that the AMES-1 PES was implemented correctly. Increasing the basis set size and modification of the atomic mass inputs did not yield any improvements, leading to the conclusion that the DVR3DRJZ input parameters had been optimised efficiently enough. To test the correct implementation of the PES, the same subroutine was adapted to work with TROVE [131] (which has been used in the current work for SO_3, see Chapters 4, 5 and 6). Vibrational band energies calculated using TROVE produced similar discrepancies. The comparisons are tabulated in Tables 3.3, and 3.4, showing vibrational bands with even and odd vibrational symmetries, respectively. In the case of the even symmetry bands, the disagreements associated with the ν_3 stretch are highlighted, and are all roughly of the order of 0.15 - 0.2 cm^{-1}, whereas other bands agree more closely. In Table 3.4 all the bands are associated with ν_3 excitation and therefore comparisons all show a similar offset.

3.2.2.2 The Improved AMES-1B PES

After ensuring that the source of the discrepancies were not related to incorrect PES implementation, basis set convergence errors or faults within the DVR3DRJZ code, the anomalies associated with AMES-1 were communicated to the original authors of the PES. An explanation for the observed disagreements was eventually found.

In a later publication [118] focused on the production of empirical line lists for various SO_2 isotopologues, Huang *et al.* discussed the lack of convergence in a particular parameter in their VTET program [179, 183]. This parameter (presumably associated with the ν_3 stretch) affects potential parts of the matrix elements, and was not set to a fully converged value for both their pre- and post-refinement calculations. The defect introduced to the potential due to this lack of convergence is corrected during the refinement procedure in VTET, and therefore their reported band origins in [35]) are actually due to an effective potential produced in VTET based on the AMES-1 PES. This means that unless other variational programs (such as DVR3D) use the same procedure as VTET then they will not necessarily produce the same band origins.

Since DVR3D does not use the same computational procedures are VTET, the AMES-1 PES could not be used to generate accurate results, and thus production on the hot line list was halted. However time was spent on making numerous modifications

Table 3.3: A comparison of even symmetry vibrational bands in cm^{-1} based on the AMES-1 PES.

Band	VTET (published) [35]	DVR3D	Difference
ν_2	517.8708	517.8692	0.0016
$2\nu_2$	1035.1121	1035.1099	0.0022
ν_1	1151.7115	1151.6996	0.0119
$3\nu_2$	1551.7447	1551.7428	0.0019
$\nu_1+\nu_2$	1666.3226	1666.3146	0.008
$4\nu_2$	2067.7815	2067.7805	0.001
$\nu_1+2\nu_2$	2180.3074	2180.3035	0.0039
$2\nu_1$	2295.8146	2295.7964	0.0182
$5\nu_2$	2583.2272	2583.2280	-0.0008
$\nu_1+3\nu_2$	2693.6859	2693.6861	-0.0002
$2\nu_3$	2713.3862	2713.2338	0.1524
$2\nu_1+\nu_2$	2807.1715	2807.1597	0.0118
$6\nu_2$	3098.0789	3098.0819	-0.003
$\nu_1+4\nu_2$	3206.4707	3206.4751	-0.0044
$\nu_2+2\nu_3$	3222.951	3222.7876	0.1634
$2\nu_1+2\nu_2$	3317.9021	3317.8965	0.0056
$3\nu_1$	3432.2839	3432.2642	0.0197
$7\nu_2$	3612.3254	3612.3312	-0.0058
$\nu_1+5\nu_2$	3718.6665	3718.6753	-0.0088
$2\nu_2+2\nu_3$	3731.9384	3731.7644	0.174
$2\nu_2+3\nu_2$	3828.0259	3828.0263	-0.0004
$\nu_1+2\nu_3$	3837.5893	3837.4214	0.1679
$3\nu_1+\nu_2$	3940.392	3940.3790	0.013
$8\nu_2$	4125.9477	4125.9569	-0.0092
$\nu_1+6\nu_2$	4230.2707	4230.2838	-0.0131
$3\nu_2+2\nu_3$	4240.3554	4240.1710	0.1844
$2\nu_1+4\nu_2$	4337.5544	4337.5607	-0.0063
$\nu_1+\nu_2+2\nu_3$	4343.7922	4343.6183	0.1739
$3\nu_1+2\nu_2$	4447.8721	4447.8657	0.0064
$4\nu_1$	4561.1041	4561.0861	0.018
$9\nu_2$	4638.9188	4638.9319	-0.0131
$\nu_1+7\nu_2$	4741.2725	4741.2898	-0.0173
$4\nu_2+2\nu_3$	4748.2031	4748.0086	0.1945
$2\nu_1+5\nu_2$	4846.4909	4846.5027	-0.0118
$\nu_1+2\nu_2+2\nu_3$	4849.4206	4849.2406	0.18
$2\nu_1+2\nu_3$	4953.5546	4953.3811	0.1735
$3\nu_1+3\nu_2$	4954.7439	4954.7437	0.0002
$4\nu_1+\nu_2$	5065.968	5065.9558	0.0122
$10\nu_2$	5151.2137	5151.2211	-0.0074

Table 3.4: A comparison of odd symmetry vibrational bands in cm^{-1} based on the AMES-1 PES.

Band	VTET (published) [35]	DVR3D	Difference
ν_3	1362.0585	1361.9235	0.1350
$\nu_2+\nu_3$	1875.7913	1875.6457	0.1456
$2\nu_2+\nu_3$	2388.9226	2388.7672	0.1554
$\nu_1+\nu_3$	2499.8906	2499.7365	0.1541
$3\nu_2+\nu_3$	2901.4657	2901.3012	0.1645
$\nu_1+\nu_2+\nu_3$	3010.3149	3010.1560	0.1589
$4\nu_2+\nu_3$	3413.4264	3413.2533	0.1731
$\nu_1+2\nu_2+\nu_3$	3520.1389	3519.9750	0.1639
$2\nu_1+\nu_3$	3629.7874	3629.6206	0.1668
$5\nu_2+\nu_3$	3924.8035	3924.6223	0.1812
$\nu_1+3\nu_2+\nu_3$	4029.3743	4029.2051	0.1692
$3\nu_3$	4054.0243	4053.9048	0.1195
$2\nu_1+\nu_2+\nu_3$	4136.9158	4136.7466	0.1692
$6\nu_2+\nu_3$	4435.5888	4435.3998	0.1890
$\nu_1+4\nu_2+\nu_3$	4538.0261	4537.8513	0.1748
$\nu_2+3\nu_3$	4559.4025	4559.2725	0.1300
$2\nu_1+2\nu_2+\nu_3$	4643.4436	4643.2712	0.1724
$3\nu_1+\nu_3$	4751.6997	4751.5258	0.1739
$7\nu_2+\nu_3$	4945.7665	4945.5702	0.1963
$\nu_1+5\nu_2+\nu_3$	5046.0924	5045.9117	0.1807
$2\nu_2+3\nu_3$	5064.2215	5064.0816	0.1399
$2\nu_1+3\nu_2+\nu_3$	5149.3807	5149.2046	0.1761
$\nu_1+3\nu_3$	5164.837	5164.7045	0.1325

to the different modules of the DVR3D suite while an improved PES was computed by Huang *et al.* These modifications, discussed in Chapter 2 were necessary to ensure rovibrational wavefunction calculations for high J were possible, as well as efficient calculations of dipole transition moments. The AMES-1 PES was used since high accuracy results were not required to optimise and test functionality of the code.

Eventually an improved PES was supplied by Huang *et al.* which maintains the same structure of 219 coefficients as the AMES-1 PES. The updated potential, AMES-1B, consists of a set of improved coefficients based on a refinement, ensuring that the PES functions correctly with variational programs other than VTET.

Table 3.5 shows the comparisons of even symmetry band origins produced by VTET and DVR3DRJZ using the new AMES-1B PES. They are in much better agreement than for the AMES-1 PES; not only is the ν_3 stretch discrepancy absent, but the average agreement has improved by one order of magnitude. The results of Huang *et*

al. using this improved potential differ slightly from their published values [35] as a result of the refinement.

Table 3.5: A comparison of even symmetry vibrational bands in cm^{-1} based on the AMES-1B PES.

Band	VTET (updated)	DVR3D	Difference
ν_2	517.8725	517.8726	-0.0001
$2\nu_2$	1035.1186	1035.1188	-0.0002
ν_1	1151.7138	1151.7143	-0.0005
$3\nu_2$	1551.7595	1551.7598	-0.0003
$\nu_1+\nu_2$	1666.3284	1666.3288	-0.0004
$4\nu_2$	2067.8084	2067.8087	-0.0003
$\nu_1+2\nu_2$	2180.3187	2180.3191	-0.0004
$2\nu_1$	2295.8152	2295.8158	-0.0006
$5\nu_2$	2583.2704	2583.2708	-0.0004
$\nu_1+3\nu_2$	2693.7053	2693.7056	-0.0003
$2\nu_3$	2713.3936	2713.3938	-0.0002
$2\nu_1+\nu_2$	2807.1739	2807.1744	-0.0005
$6\nu_2$	3098.1428	3098.1432	-0.0004
$\nu_1+4\nu_2$	3206.5009	3206.5012	-0.0003
$\nu_2+2\nu_3$	3222.9523	3222.9526	-0.0003
$2\nu_1+2\nu_2$	3317.9078	3317.9082	-0.0004
$3\nu_1$	3432.2724	3432.2729	-0.0005
$7\nu_2$	3612.4145	3612.4150	-0.0005
$\nu_1+5\nu_2$	3718.7109	3718.7111	-0.0002
$2\nu_2+2\nu_3$	3731.937	3731.9373	-0.0003
$2\nu_2+3\nu_2$	3828.0367	3828.0370	-0.0003
$\nu_1+2\nu_3$	3837.6154	3837.6161	-0.0007
$3\nu_1+\nu_2$	3940.3781	3940.3786	-0.0005
$8\nu_2$	4126.0668	4126.0673	-0.0005
$\nu_1+6\nu_2$	4230.3325	4230.3327	-0.0002
$3\nu_2+2\nu_3$	4240.3549	4240.3553	-0.0004
$2\nu_1+4\nu_2$	4337.5726	4337.5729	-0.0003
$\nu_1+\nu_2+2\nu_3$	4343.8153	4343.8158	-0.0005
$3\nu_1+2\nu_2$	4447.8567	4447.8572	-0.0005
$4\nu_1$	4561.0634	4561.0638	-0.0004
$9\nu_2$	4639.0726	4639.0731	-0.0005
$\nu_1+7\nu_2$	4741.3554	4741.3556	-0.0002
$4\nu_2+2\nu_3$	4748.2069	4748.2074	-0.0005
$2\nu_1+5\nu_2$	4846.519	4846.5192	-0.0002
$\nu_1+2\nu_2+2\nu_3$	4849.4433	4849.4438	-0.0005
$2\nu_1+2\nu_3$	4953.5971	4953.5978	-0.0007
$3\nu_1+3\nu_2$	4954.7285	4954.7289	-0.0004
$4\nu_1+\nu_2$	5065.9188	5065.9192	-0.0004
$10\nu_2$	5151.3969	5151.3974	-0.0005

3.2.2.3 The AMES-1 DMS

The dipole moment surface used in the calculations of the current project was also supplied by Huang *et al.* [35]. A total of 3638 dipoles were calculated using a CCSD(T)/aug-cc-pV(Q+d)Z level of theory (a better description of the diffuse part of the wavefunction is necessary for the calculation of electrical properties, hence the use of the aug-cc-pV(Q+d)Z basis set). These dipoles have been fitted to an expansion formula containing a total of 969 coefficients. Both the formula and expansion coefficients are encapsulated into a subroutine and coefficient file, to be implemented into the DIPOLE3 code.

In order to be used by DIPOLE3 the subroutine must be adapted to return the value of the dipole *DIPC* as a function of the input arguments R_1, R_2, and *XCOS*, as in the previous case for the potential V, as well as NU, which represents the $NUth$ component of the dipole, where $NU = 0$ corresponds to μ_z and $NU = 1$ corresponds to μ_x. In the bisector embedding only the z and x components of the dipole are non-zero, and for a given molecular configuration in R_1, R_2, and *XCOS*, the subroutine is called twice to compute the dipole across both these components, in order to calculate the strength of a particular transition [138]. The orientation of the dipole must therefore be carefully considered when implementing the subroutine due to the direction-dependence of μ_z and μ_x.

3.2.3 Computing the Line List

All calculations were carried out on the High Performance Computing Service Darwin cluster, located at Cambridge. Each job from DVR3DRJZ, ROTLEV3B and DIPOLE3 is submitted to a single computing node consisting of two 2.60 GHz 8-core Intel Sandy Bridge E5-2670 processors, therefore making use of a total of 16 CPUs each through OpenMP parallelisation of the various BLAS routines in each module. A maximum of 36 hours and 64 Gb of RAM are available for each calculation on a node.

Submitting jobs to the compute nodes is subjected to a fair-usage queuing policy, where job priority is based on the amount of requested resources per job. The DVR3DRJZ runs generally do not require more than 2 hours of wall clock time. The most computationally demanding parts of the line list calculation are in ROTLEV3B for the diagonalisation of the Hamiltonian matrices, where wall clock time increases

by a factor of N^3 with increasing J, where N is the matrix dimension. For each job in ROTLEV3B the maximum memory allocation is requested, which ensures all diagonalisations of the matrices are possible using the LAPACK DSYEV routine [186].

3.2.3.1 Rovibrational Wavefunctions

The input parameters required for all DVR3DRJZ and ROTLEV3B calculations are summarised in Table 3.2. To compute rovibrational wavefunctions for a particular value of J it is necessary to run the DVR3DRJZ module twice to produce both even and odd symmetry vibrational wavefunctions files. Each of these files is then used in a ROTLEV3B run to produced rovibrational wavefunctions of one of four vibrational symmetry-rotational parity combinations, ee, eo, oo, or oe (e = even, o = odd). The required combinations depend on whether J is even or odd. Table 2.1 sums up the physically allowed symmetry combinations for rovibrational states.

The desired energies for use in calculating the line list are chosen to be below 15 000 cm^{-1}. Some code was added to the ROTLEV3B module to display the amount and sums of energies below 8000, 9000, 10 000, 11 000, 12 000, 13 000, 14 000, and 15 000 cm^{-1}. Initially a number of ROTLEV3B runs were performed at various values of J in increasing steps of 10 without limiting the number of eigenvalues displayed by the Hamiltonian diagonalisation. This gave an idea of the number of energies as a function of J were under 15,000 cm^{-1} and therefore it was possible to fit a functional form to determine the amount of energies that needed to be saved to disk for all J. This also allowed for automation of the input files for ROTLEV3B. The specified number of energies required for each run was slightly overestimated to ensure the entire range was covered.

Figure 3.1 shows both the dimension of the hamiltonian matrix ($IBASS = n(J+1)$) and number of energies below the 15 000 cm^{-1} threshold as a function of J (for all rotational parities). At J = 124 the dimension of the hamiltonian is reduced due to a computational bottleneck in the DSYEV eigensolver; up to this point, all matrix diagonalizations are carried out over 16 CPUs via OpenMP parallelization, however at J = 124, with $NBASS$ = 90 625 (for even rotational parity), it is impossible to complete this within the 36 wall clock limit imposed on the Darwin cluster, as the computational time increases roughly with N^3, where N is equal to the matrix dimension.

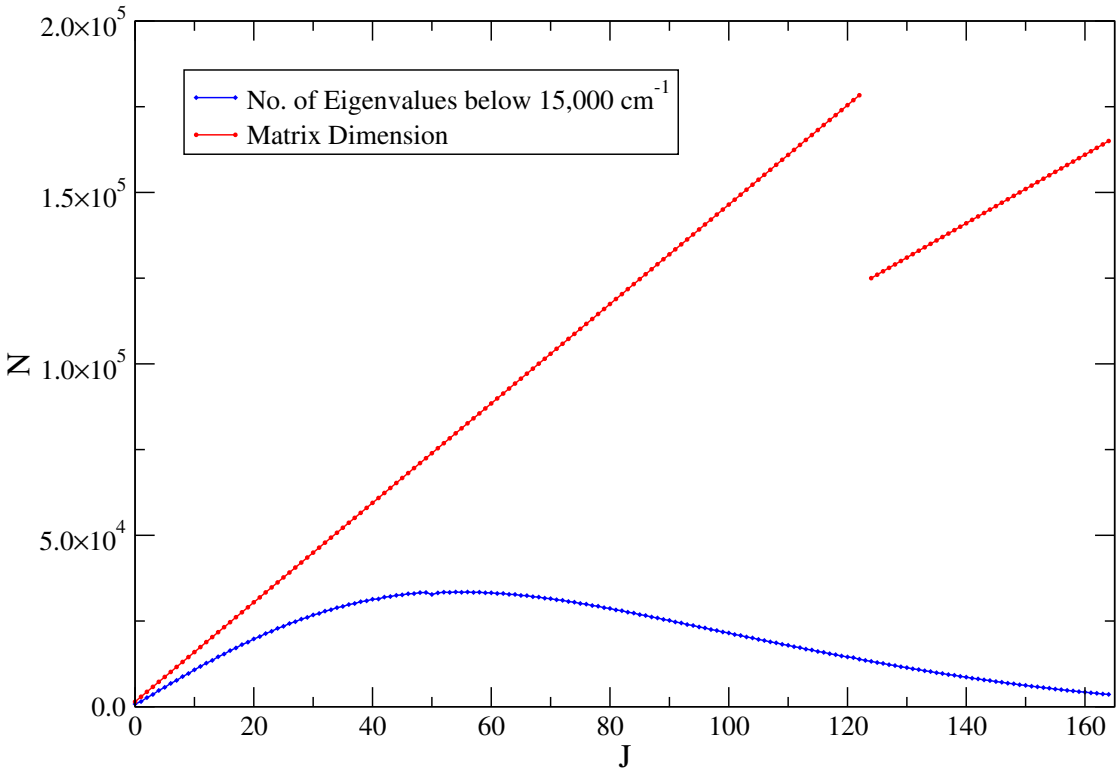

Figure 3.1: Size of Hamiltonian matrix (for both even and odd rotational parities combined) and number of eigenvalues below 15 000 cm^{-1} as a function of increasing J. The discontinuity in the matrix dimension reflects the decrease in the value of *NBASS* for calculations above $J = 123$. There is also a small discontinuity in the blue curve at $J = 50$ due to a defect in the DVR3D calculations. Both of these are referred to in the main text.

Reducing the value of n in $IBASS = n(J+1)$ to 500 was found to have essentially no effect on the convergence of rovibrational eigenvalues produced at $J = 124$. At this point there exists a total of 14 523 eigenvalues below 15 000 $cm^{-1}$for both rotational parities combined. This constitutes roughly 8% of the total combined matrix dimension of 175 450 for $n = 725$ and 12% of 121 000 for $n = 500$. The value $n = 725$ was originally obtained for convergence of energies at $J = 60$, where the number of eigenvalues below 15 000 cm^{-1} accounts for 38% of the combined matrix dimension of 88 450. The higher energies at $J = 60$ are much more sensitive to the value of n due to the way the basis functions provided by DVR3DRJZ are distributed, whereas for $J \geq 120$ the energies below the 15 000 cm^{-1} threshold are already easily converged at lower values of n.

During the production of the line list a defect in the AMES-1B PES was identified which severely affected the computed high J eigenvalues; at small bond angles, the PES

appears to become negative. During a particular stage in the DVR3DRJZ calculation the angular functions corresponding to the lowest $MAX3D$ = 3052 eigenvalues of the 1D angular Hamiltonian are selected for use in the construction of the 3D hamiltonian. In the DVR, this corresponds to the selection of functions located over a range of DVR points. Due to the presence of negative eigenvalues associated with the potential defect, the sampling of functions was carried out incorrectly, where the lowest negative eigenvalue was treated as the 'true' lowest value.

Figure 3.2 shows that the angular functions selected (associated with the lowest eigenvalues) tend to come from around the same region in the 1-D DVR grid space, corresponding to medium-high angle in the θ coordinate. The small bump at γ = 65 - 70 shows an example of function sampling which includes functions at grid points associated with high θ via the AMES-1B PES. The DVR3DRJZ program ensures that these functions are discarded as they are assumed to be non-physical. However with increasing J (and therefore, k), it becomes ambiguous as to when such functions should be discarded.

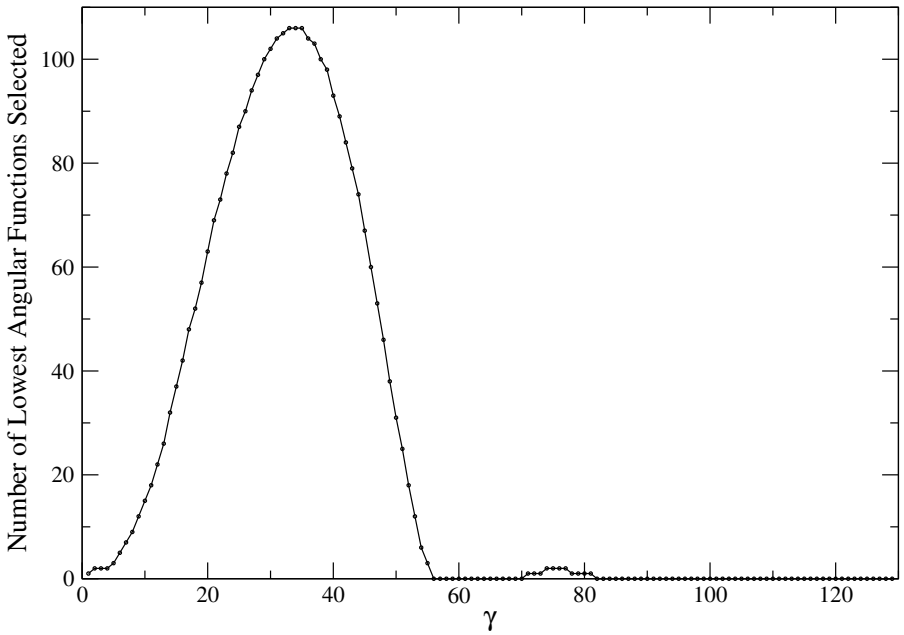

Figure 3.2: Plot of the number of lowest angular functions selected as a function of γ, the angular grid point index, where γ_{max} = 130.

Although DVR3DRJZ discards the unphysical functions, the distribution of low 1D angular energies is incomplete in that it is less by the same amount as there are discarded functions, i.e. N = 3052 - $N_{discarded}$. This affects the ROTLEV3B basis

functions provided by the DVR3DRJZ module, and manifests as the high energy functions being slightly lower. A small amendment to the DVR3DRJZ code allowed for the selection of appropriate functions by ignoring those above a certain DVR grid point, and distributed the 3052 lowest functions across those that remained. This amendment was essential for the high J calculations where higher rovibrational energies start to dominate.

This version of the code was used for all $J \geq 50$. For $J < 50$ this defect only affected very high rovibrational energies computed in ROTLEV3B. This can be seen via the blue curve in Figure 3.1 at the small $J = 50$ discontinuity, where the number of eigenvalues below 15 000 cm^{-1} is slightly higher for $J < 50$ due to the defected PES.

A total of 3 255 954 energy levels below 15 000 cm^{-1} have been calculated for all $J \leq 165$.

3.2.3.2 Transition Moments

For each value of J there are two ROTLEV3B-produced rovibrational wavefunction files to be used for DIPOLE3 calculations. The code added to ROTLEV3B for determining the number of energies below a threshold is also useful for the DIPOLE3 runs; the input allows the number of bra (lower) energies and the number of ket (upper) energies to be specified for the calculation. For the current work all energies below 8000 cm^{-1} are considered for the bra inputs to DIPOLE3, and all energies below 15 000 cm^{-1} are considered for ket inputs. The code added to ROTLEV3B gives the required numbers, and this greatly speeds up the DIPOLE3 runs as superfluous energies are not considered in the dipole transition moment calculations.

The DIPOLE3 runs also make use of the maximum memory per node, which enables records from the wavefunction files to be stored as arrays to speed up the computation. This was a new modification introduced for the current SO_2 project, discussed in Chapter 2.

Due to the huge sizes of the wavefunction files the danger of overloading the hard disk space was an issue. For the most energy-dense values of J the files lie above or around 0.5 Tb, so computing the wavefunction files for all values of J at once was untenable. Therefore a procedure was adopted of computing wavefunctions in batches of 10 J's, i.e. all files for $J = 0 \rightarrow 9$, $10 \rightarrow 19$, $20 \rightarrow 29$, etc. After each batch is

computed, relevant DIPOLE3 calculations are performed using these files, after which the wavefunction files are discarded, leaving disk space open for the following batch.

In total, there are 1.3 billion transitions computed for $^{32}S^{16}O_2$ between all energies with $J \leq 165$, $E'' \leq 8000$ cm^{-1} $E' \leq 15\ 000$ cm^{-1}. Figure 3.3 shows the number of transitions between J pairs as a function of J''.

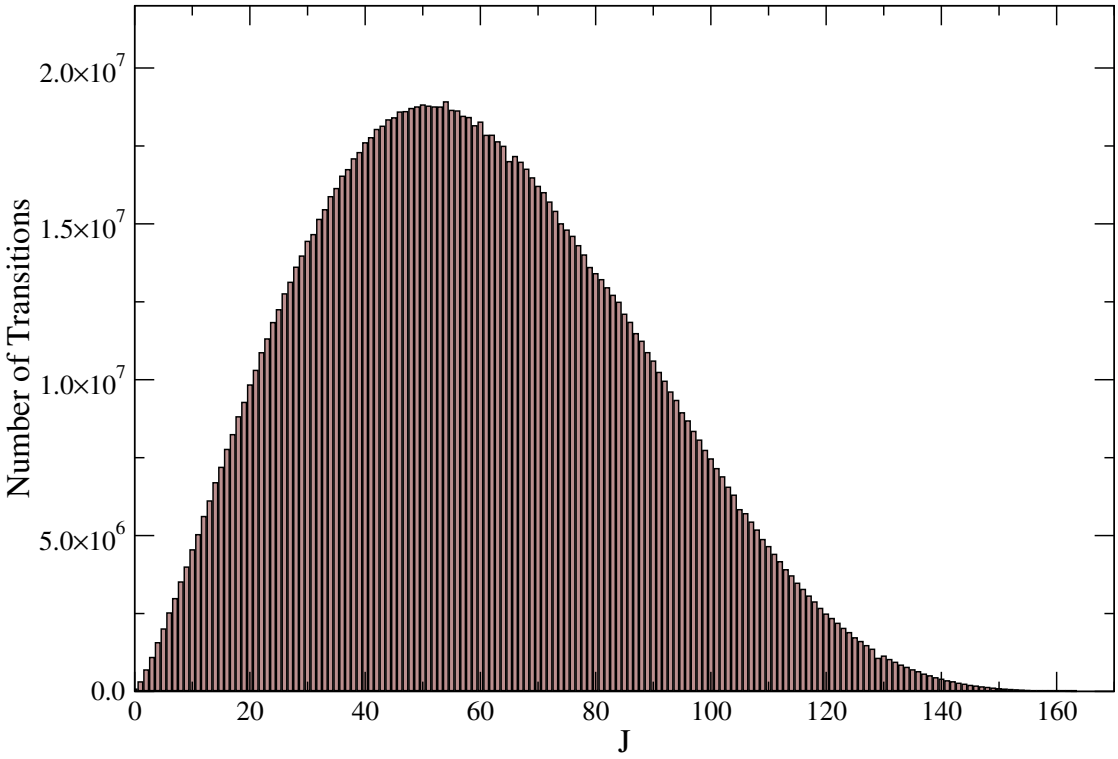

Figure 3.3: Plot of the number of transitions calculated from energy levels with rovibrational quantum number J.

3.3 Temperature Dependence and Completeness of the Line List

The current work for SO_2 seeks to expand upon the work of Huang *et al.* by the production of a line list suitable for the modelling of spectra above room-temperature. As described in Chapter 2, the convergence of the partition function, Q, at the given temperature, T, is necessary in order for computed line list to be applicable at this temperature. The computation of wavefunctions corresponding to energy levels greater than the cut-offs imposed by Huang *et al.* ($J \leq 80$, E$\leq$ 13 168 cm^{-1}) seeks to achieve convergence for $T > 296$ K.

The value of the partition function at T = 296 K, computed using all energy levels

satisfying $J \leq 165$ and $E \leq 15\,000$ cm^{-1}, is 6337.131. With a cut-off of $J \leq 80$, as used in the work of Huang *et al.*, the value for the same temperature is computed as 6336.803, which is in excellent agreement with their calculated value of 6336.789.

Figure 3.4 shows the partition function values as a function of a J cut-off for a range of temperatures. The highest value of J considered, $J = 165$, defines the last point where the lowest energy is less than 8000 cm^{-1}, which is used as the maximum value of lower energy states in DIPOLE3 calculations. As can be seen from this figure, the partition function is well converged for $J = 165$ at all temperatures. Table 3.6 shows the final values of Q obtained for each temperature, quoted alongside their degree of convergence.

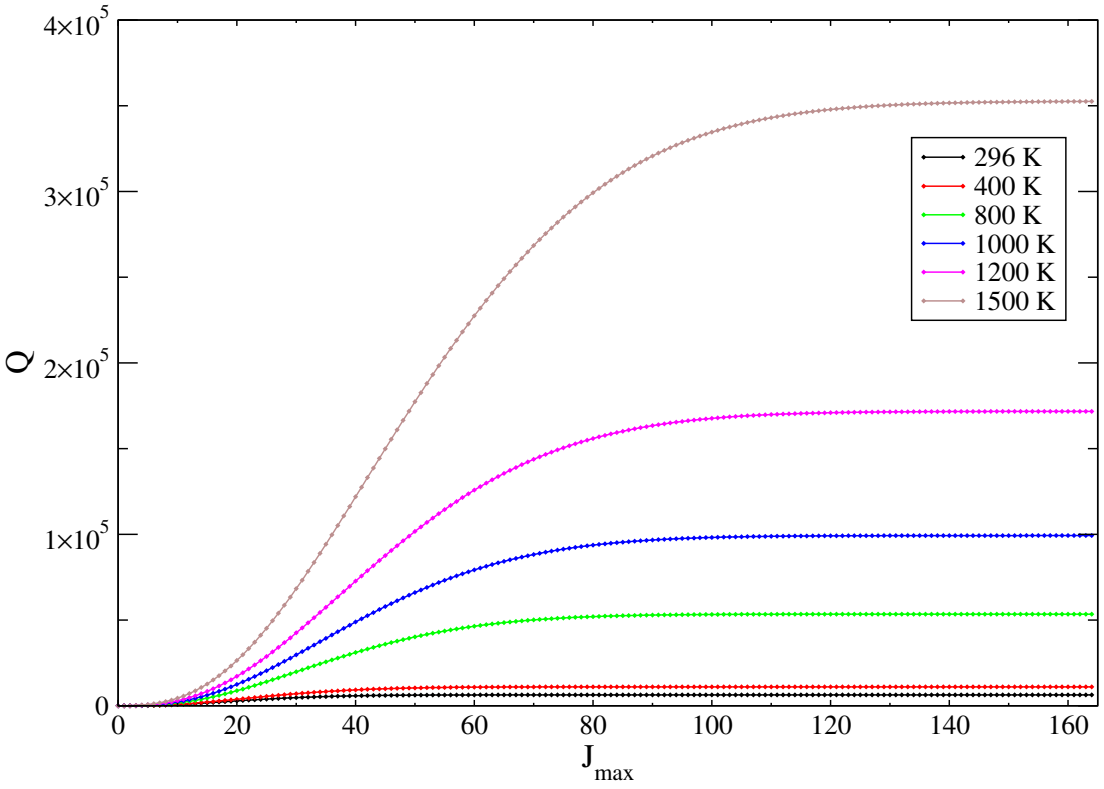

Figure 3.4: Convergence of partition function at different temperatures as a function of J_{max}.

The J-dependent convergence of Q gives a good indication of the completeness of the computed energy levels with respect to their significance at each temperature. However in order to ascertain the reliability of the line list for increasing temperatures it is more pertinent to observe the convergence of Q as a function of energy cut-off; this is illustrated in Figure 3.5.

The importance of this lies in the fact that the computed line list in the current work

Table 3.6: Values of the partition function, Q, for different temperatures, T. The degree of convergence is specified by $Q_{J_{165}} - Q_{J_{164}}/Q_{J_{165}} \times 100$.

T (K)	Q	Degree of Convergence (%)
296	6337.131	$< 10^{-12}$
400	11040.245	1.8×10^{-12}
800	53449.248	2.9×10^{-6}
1000	99288.461	4.6×10^{-5}
1200	171740.274	2.8×10^{-4}
1500	352549.889	1.6×10^{-3}

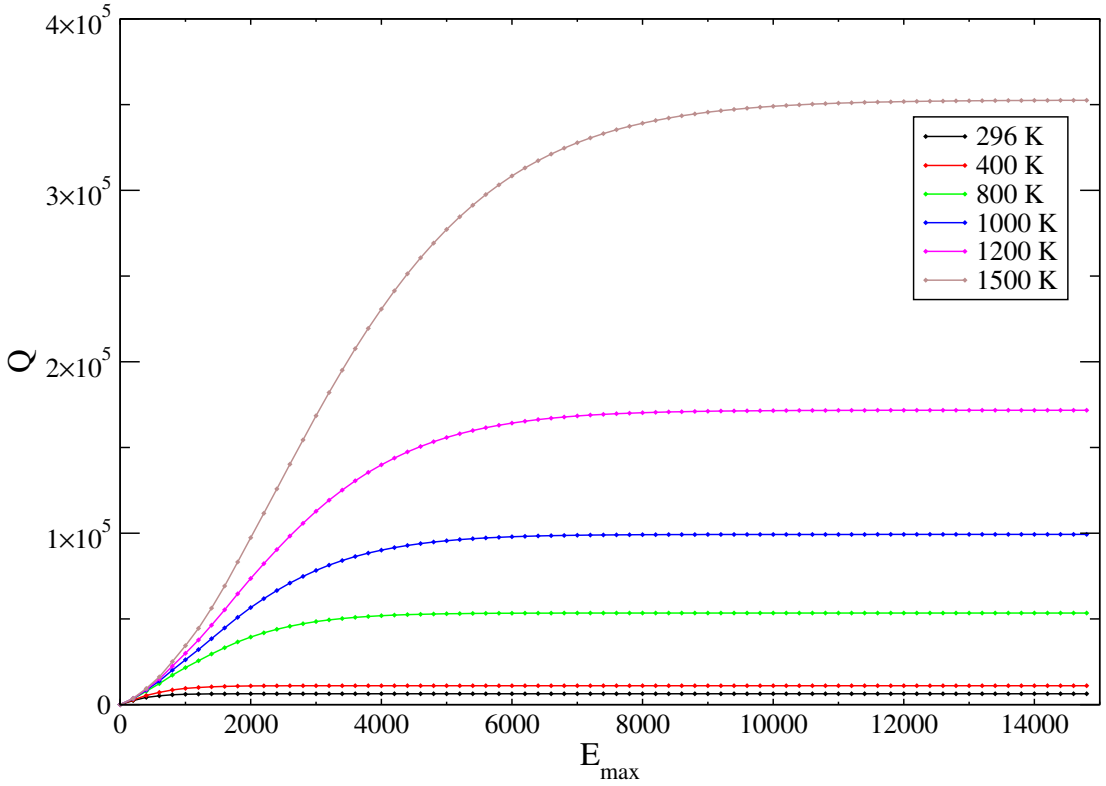

Figure 3.5: Convergence of partition function at different temperatures as a function of E_{max} (cm^{-1}).

only considers transitions from energy levels below 8000 cm^{-1}. Since the physical interpretation of an energy level's contribution to Q is the probability of it's occupancy, the completeness of the line list can only be guaranteed if all transitions from states with non-negligible population are computed. In other words, the line list may only be considered 100 % complete if Q is converged when summing over all $E \leq 8000$ cm^{-1}.

Figure 3.5 shows that, at a cut-off of 8000 cm^{-1}, the partition function is not fully converged for $T = 1500$ K. Despite computing all rovibrational levels below 8000 cm^{-1} ($J \leq 165$), and all transitions from these states to states with $E \leq 15\,000$ cm^{-1},

there is still a minor contribution from energies above this cut-off to the partition sum, corresponding to all values of J. However, the neglected transitions are not expected to provide significant addition to the overall intensity.

The completeness of the line list may be quantified by considering the ratio of the partition function at the 8000 cm^{-1} cut-off and the total partition function, Q_{Total}, which takes into account all computed energies. Figure 3.6 shows this ratio as a function of temperature.

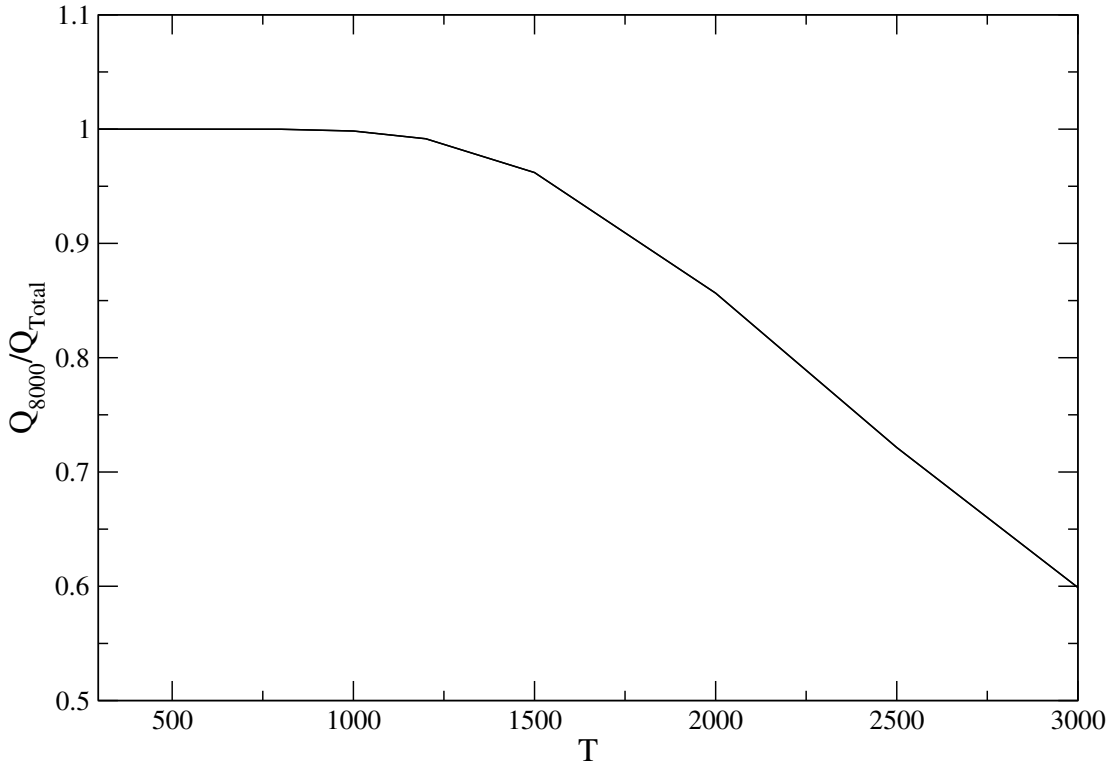

Figure 3.6: Ratios of Q_{8000} to the assumed converged values Q_{Total} as a function of temperature.

For $T \leq 1500$ K the line list is over 96 % complete. As can be seen from Figure 3.6 the level of completeness decreases with increasing temperature; at 2000 K the ratio falls to 86 %, and as low as 33 % for 5000 K. The assumption made here is that Q_{Total} is equal to the 'true' value of the partition function; in reality, this depends on the ability of the AMES-1B PES to accurately reproduce the energies that become important at these higher temperatures. Nevertheless, it is interesting to note that at 3000 – 5000 K the partition function is still converged to within 0.1 % when all computed energy levels are taken into consideration, though the 8000 cm^{-1} cut-off means that the line list becomes less than 50 % complete at these temperatures, as the contribution to the

Table 3.7: Extract from the state file for SO_2. Quantum numbers are defined in Table 3.8.

i	$\tilde{E}$	g	J	p	ν_1	ν_2	ν_3	K_a	K_c
1	0.000000	1	0	0	0	0	0	0	0
2	517.872609	1	0	0	0	1	0	0	0
3	1035.118794	1	0	0	0	2	0	0	0
4	1151.714304	1	0	0	1	0	0	0	0
5	1551.759779	1	0	0	0	3	0	0	0
6	1666.328818	1	0	0	1	1	0	0	0
7	2067.808741	1	0	0	0	4	0	0	0
8	2180.319086	1	0	0	1	2	0	0	0
9	2295.815835	1	0	0	2	0	0	0	0
10	2583.270841	1	0	0	0	5	0	0	0
11	2693.705600	1	0	0	0	4	0	0	0
12	2713.393783	1	0	0	0	0	2	0	0
13	2807.174418	1	0	0	2	1	0	0	0
14	3098.143224	1	0	0	0	6	0	0	0
15	3206.501197	1	0	0	0	5	0	0	0
16	3222.952550	1	0	0	0	1	2	0	0
17	3317.908237	1	0	0	1	3	0	0	0
18	3432.272904	1	0	0	3	0	0	0	0
19	3612.415017	1	0	0	0	7	0	0	0
20	3718.711074	1	0	0	0	6	0	0	0

partition function from energy levels higher than 10 000 cm^{-1} starts to dominate.

The line list is presented in the ExoMol format [175] described in Chapter 2, i.e. via a 'states', containing a list of energy levels, and a 'transition' file which lists upper and lower energies via their index from the 'states' file, and associated Einstein A coefficient. Table 3.7 gives a portion of the SO_2 states file; the quantum numbers are described in Table 3.8. As DVR3D does not provide approximate quantum numbers, K_a, K_c and the vibrational labels ν_1, ν_2 and ν_3, these have been taken from the calculations of [35], where possible, by matching J, parity and energy (see Section 3.4); these quantum numbers are approximate and may be updated in future as better estimates become available. Table 3.9 gives a portion of the transitions file.

3.4 Rovibrational Energy Levels & Intensity Simulations

The validation both the CCSD(T)/cc-pVQZ-DK Ames-1 PES and of the CCSD(T)/aug-cc-pV(Q+d)Z DMS for room-temperature simulations is carried out by Huang *et al.*

Table 3.8: Quantum numbers used in labelling energy states.

Quantum Number Descriptions
i: State counting number.
$\tilde{E}$: State energy in cm^{-1}.
g: State degeneracy.
J: Total angular momentum
p: Total parity given by $(-1)^{J+p}$.
v_1: Symmetric stretch quantum number.
v_2: Bending quantum number.
v_3: Asymmetric stretch quantum number.
K_a: Asymmetric top quantum number.
K_c: Asymmetric top quantum number.

Table 3.9: Extract from the transitions file for SO_2: f is the upper state counting number, i is the lower state counting number, and A_{fi} is Einstein-A coefficient in s^{-1}.

f	i	A
679	63	1.9408E-13
36	632	5.6747E-13
42	643	1.7869E-11
635	38	1.1554E-11
54	662	3.6097E-11
646	44	1.9333E-08
660	52	2.5948E-08
738	98	3.4273E-06
688	69	3.4316E-06
47	650	1.4537E-11
648	45	3.4352E-06
711	82	3.5730E-06
665	55	3.5751E-06
716	85	3.4635E-06
670	58	3.4664E-06
635	37	3.4690E-06
611	23	3.4701E-06
595	12	3.4709E-06
734	95	3.7253E-06
684	66	3.7257E-06

[35], however it is useful to re-assess their accuracy within the context of the current work, particularly since the Ames-1B PES is employed in the DVR3D calculations. The correct implementation of both into the DVR3D suite must be tested, and any major discrepancies obtained via the differing methodologies should be addressed.

The purpose of the current work is to produce a hot line list by the extension of rovibrational energy and intensity calculations, and is primarily concerned with the

adaptation of the DVR3D program suite in order to achieve this aim. It should therefore be noted that substantial improvement is not necessarily anticipated when comparing the current work to experimental data; there has been no focus on the improvement in the ability of the PES and the DMS to accurately replicate experimental energies and intensities. The convergence of DVR3D wavefunctions is expected to at least allow for comparable data analyses to those already carried out by Huang *et al.*

3.4.1 Comparison with Ames-296K

The energy level values published by Huang *et al.* in Ref. [35] are derived from a different PES to that used in the current work (see above). It is difficult to compare energy levels, particularly at higher energies due to ambiguities arising from possible double matching based on quantum numbers. However the self-contained "complete" nature of both line lists including transition intensities allows for suitable comparisons to be made.

In an approach similar to one made in Ref. [35], transitions between both data sets have been compared in order to asses both the agreement between energy levels/line positions, and line intensities, for T = 296 K. Huang *et al.* made use of the rigorously defined quantum numbers J and the rotational parity in order to match their lines with corresponding transitions in HITRAN, along with a restriction on the maximum deviation of wavenumbers for line positions and energy levels. These rules were originally chosen by Huang *et al.* to avoid discrepancies observed due to large experimental uncertainties and disagreements of quantum number assignments between observed and computed values of line positions are intensities [35]. Since the DVR3D method does not readily provide description of the vibrational bands or K_a and K_c quantum numbers, the criteria of parity matching is well-suited for this comparison.

The criteria used here for line matching are given as: Obs. - Calc. residuals for E_i and ν are ≤ 0.1 cm^{-1}, and the quantum numbers J_f, J_i, and the parity (defined by K_a + n_{ν_3}) agree. In addition, intensity comparisons are screened via the symmetric residual [35] $\delta(\mathrm{I})\% = 50\% \times (\mathrm{I}_{DVR3D}/\mathrm{I}_{Ames-296K} - \mathrm{I}_{Ames-296K}/\mathrm{I}_{DVR3D})$, where the restriction of $\delta(\mathrm{I})\% \leq 10\%$ is used. The inclusion of intensities facilitates the comparison with the imposed restriction allowing for lines to be matched, as well as the energy states. Based on these criteria a total of 4 031 063 lines are matched to the 4 196 337 lines

in the Ames-296K line list with an intensity greater than 10^{-30} cm molecule^{-1}. This translates to roughly 96% of lines matched. In reality it is difficult to rule out double matches in such a comparison, and therefore in these cases the match with the minimum deviations is kept. In general, the largest deviations in line positions are due to higher energy disagreements for energy levels. To give an indication of the quality of the matches, 94% of matches are ≤ 0.05 cm^{-1} in line position, and 97% are ≤ 0.1 cm^{-1}. The standard deviation in line positions is 0.043 cm^{-1}, and for transition intensities the standard deviation in the symmetric residual is given as $\sigma(\delta \mathrm{I}) = 1.63\%$.

Figure 3.7 shows an overview comparison of simulated spectra at 296 K using the line list generated in the current work, using the DVR3D method in conjunction with the Ames-1B PES and DMS, and the VTET-produced Ames-296K line list provided in the supplementary material of Ref. [35]. The lowest intensity considered is 10^{-30} cm molecule^{-1}. The DVR3D method reproduces excellently the intensity features generated by VTET at this temperature, indicating an effective implementation of the PES and DMS into the DVR3D code, and validating the adaptations made to each program module in order to make the calculations possible.

Minor differences begin to appear as ν approaches 8000 cm^{-1}, where the hot line list provides a very small contribution to the overall intensity in this spectral region, of the order 10^{-29} cm molecule^{-1}, as a consequence of the consideration of higher energies than in the Ames-296K line list. Such intensity contributions are present across the entire $0 < \nu \leq 8000$ cm^{-1} region when $J > 80$ energies are taken into consideration, but are negligible in comparison to higher intensity transitions which are already provided by Ames-296K.

3.4.2 Comparison with HITRAN

There are 72 459 lines for $^{32}S^{16}O_2$ included in the HITRAN2012 database [75], which include rovibrational energies up to and including $J = 99$. In order to quantitatively compare energy levels and absolute intensities a similar approach was adopted to that of Huang *et al.* in Ref. [35].

3.4.2.1 Rovibrational Energy Levels

In order to compare energy levels, the HITRAN transitions are transformed into a list of levels labelled by their appropriate upper and lower state quantum numbers; ener-

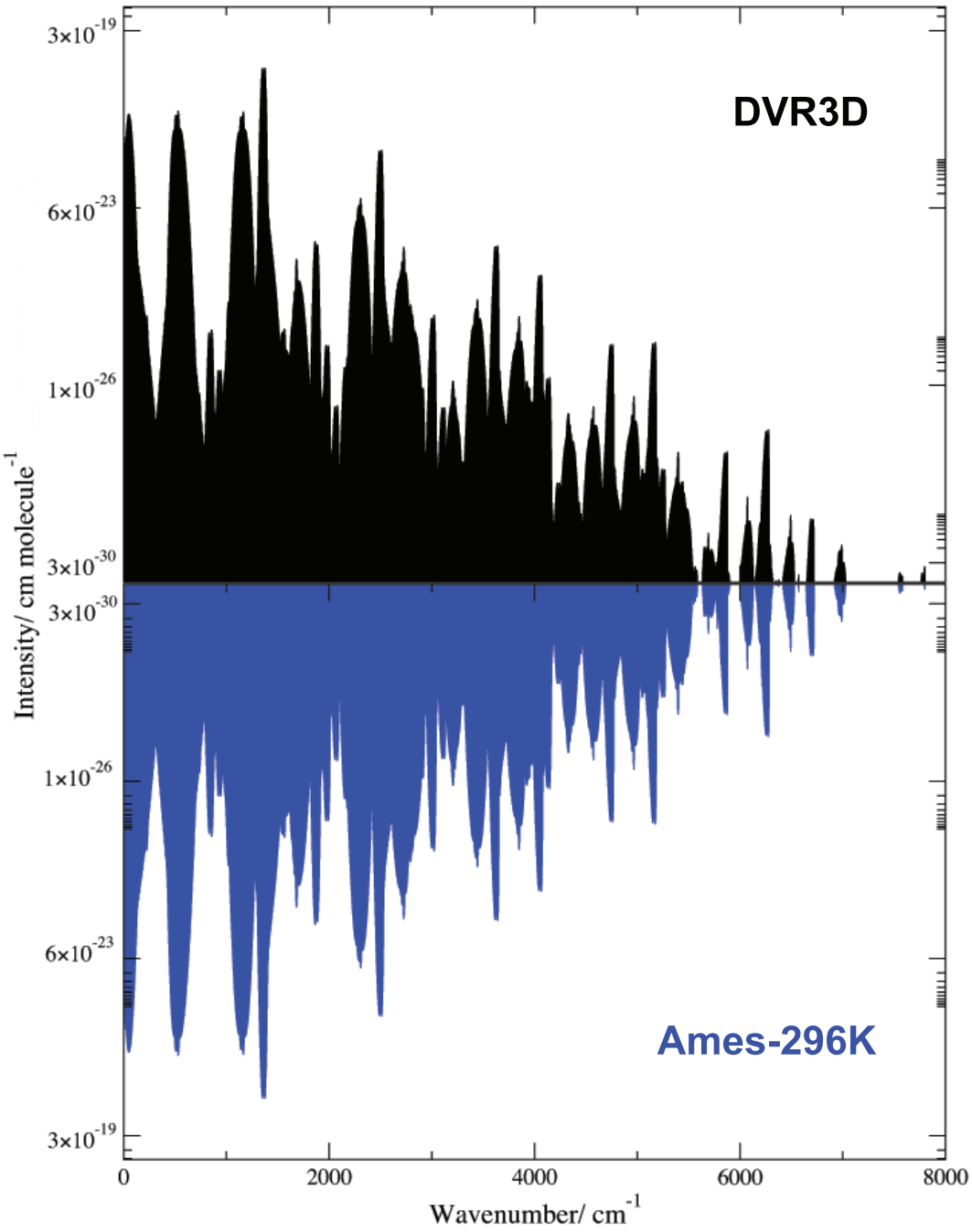

Figure 3.7: Comparison of simulated spectra at T = 296 K, between line lists generated in the current work (black), and by Huang *et al.* [35] (blue).

gies are obtained from the usual lower energy column, E'', and upper energies are also obtained via $E'' + \nu$. Any duplication from the combination difference method is removed, and energies are only kept if the HITRAN error code for line position satisfies the condition $ierr \geq 4$, ensuring all line position uncertainties are under 1×10^{-3} cm^{-1}. For this reason, the $\nu_1 + \nu_3$, $\nu_1 + \nu_2 + \nu_3$ and $3\nu_3$ bands are excluded from the current comparison, as in Ref. [35]. This leaves a total of 13 507 rovibrational levels across 10

vibrational bands available for comparison.

Energies are matched in a similar fashion to that in the comparison with Ames-296K, based on the following criteria: Obs. - Calc. residuals for E_i are ≤ 0.2 cm^{-1}, and the quantum numbers J_f, J_i, and the parity (defined by $K_a + n_{v_3}$) agree. The results are displayed in Table 3.10.

Table 3.10: Comparisons of rovibrational energy levels between available HITRAN data [75] and corresponding data calculated using DVR3D.

v_1 v_2 v_3	E_{min}	E_{max}	J_{min}	J_{max}	K_a^{min}	K_a^{max}	No.	Δ_{max}	Δ_{RMS}
0 0 0	1.908	4062.964	1	99	0	35	2774	0.092	0.014
0 0 1	1362.696	4085.476	1	90	0	33	2023	0.092	0.019
0 0 2	2713.383	4436.384	0	76	0	23	1097	0.085	0.013
0 1 0	517.872	3775.703	0	99	0	29	2287	0.084	0.016
0 2 0	1035.126	2296.506	0	62	0	20	894	0.073	0.010
0 3 0	1553.654	2237.936	0	45	0	17	502	0.070	0.016
1 0 0	1151.713	3458.565	0	88	0	31	1706	0.097	0.016
1 1 0	1666.335	3080.042	0	45	0	21	757	0.080	0.007
0 1 1	1876.432	3964.388	1	70	0	25	1424	0.087	0.017
1 3 0	2955.938	3789.613	11	52	11	11	43	0.075	0.057
Total	1.908	4436.384	0	99	0	35	13507	0.097	0.016

Overall the agreement with the corresponding comparison in Ref. [35] is fairly consistent. There are some minor deviations in Δ_{max}, though the values of Δ_{RMS} are comparable. However it is not entirely meaningful to compare to the values reported by Huang *et al.* It is likely that these deviations are largely due to the use of the Ames-1B PES in the DVR3D calculations. It is unclear whether there is any contribution based on the inclusion of energies above $J = 80$, and computational deviations introduced by the differing methods.

3.4.2.2 Absolute Intensities

In keeping with the replication of comparison analyses, HITRAN band positions and intensities are compared to the data produced in the current work, again in a similar fashion to Huang *et al.* In Ref. [35] all 13 HITRAN bands are compared (despite three of these being excluded from their energy level comparisons). In their comparison all transitions associated with $2\nu_3$ and $K_a = 11$ levels were excluded due to a resonance of the band with $\nu_1 + 3\nu_2$; the same exclusion has been applied here.

A total of 70 830 transitions are available for comparison here, taking into account

those corresponding to energy levels with $J > 80$. The matching criteria are identical to that outlined above for energy levels, with the addition that the Obs. - Calc. residuals for ν also satisfy $\leq 0.2\ \mathrm{cm}^{-1}$. The algorithm used is prone to double-matching, leading to comparisons which may be reasonable in wavenumber residuals but not in intensity deviations. In these instances, the intensity comparisons are screened via the symmetric residual [35] $\delta(\mathrm{I})\% = 50\% \times (\mathrm{I}_{DVR3D}/\mathrm{I}_{HITRAN} - \mathrm{I}_{HITRAN}/\mathrm{I}_{DVR3D})$, where the best match is found where this value is at a minimum. These criteria have been able to match all available lines, with the exception of the $001 \leftarrow 000$ which matches 5686 out of 5721 lines. The explanation for this appears to be due to systematic errors in HITRAN for energies with $K_a \geq 33$ [35, 86]. Table 3.11 shows a statistical summary of the band comparisons in a similar fashion to that of Huang *et al.*

The standard deviations in line position, $\sigma(\Delta\nu)$, and line intensity, $\sigma(\delta\mathrm{I})$, are in fairly good agreement with those of Huang *et al.* despite the used of a different PES, and the inclusion of energies with $J > 80$, as mentioned above in the case of the energy level comparisons. The differences in minimum, maximum, and average values may be attributed to these, though the tighter restriction on the line intensity matching algorithm used in this work may also contribute.

3.4.3 Comparison with Recent High-Temperature Measurements

The reproducibility of the Ames-296K data and the comparison with HITRAN serve as an important indicator of the correct implementation of the DVR3D method in approaching the SO_2 problem. However the main purpose of the current project is in the calculation of a line list which is appropriate for high-temperature modelling.

There is limited spectral data for SO_2 available at elevated temperatures, and is either not applicable to the spectral region of interest or consists of remote observational data requiring sophisticated, bespoke atmospheric modelling to reproduce in conjunction with a line list [3,42]. However a few measurements of cross section data have been made for hot SO_2 spectra in the laboratory, carried out by Fateev *et al.* [42, 43, 187], the group motivating the work of this thesis (see Chapter 1). This data is not publicly available at the time of writing, but has been obtained collaboratively and via private communication, and is included here for a preliminary comparison with simulated cross sections obtained from the hot line list.

Table 3.11: Statistical summary of comparisons between 13 HITRAN bands and corresponding bands produced in the current work. Transition frequencies ν are given in cm^{-1} and intensities are in cm $molecule^{-1}$.

Band	J_f	K_a	ν_{min}	ν_{max}	No.	$\Delta\nu_{max}$	$\Delta\nu_{AVG}$	$\sigma(\Delta\nu)$	Sum I_{HITRAN}	Sum I_{DVR3D}	δI_{min}	δI_{max}	δI_{AVG}	$\sigma(\delta I)$
000 ← 000	0 – 99	0 – 35	0.017	265.860	13725	0.048	0.004	0.007	2.21×10^{-18}	2.39×10^{-18}	-5.3%	71.0%	14.8%	11.0%
010 ← 010	0 – 99	0 – 29	0.029	201.901	9215	0.041	0.004	0.006	1.78×10^{-19}	1.93×10^{-19}	0.6%	40.6%	10.6%	5.8%
010 ← 000	0 – 70	0 – 26	436.589	645.556	5914	0.030	0.006	0.005	3.71×10^{-18}	3.84×10^{-18}	-48.7%	38.9%	2.5%	18.8%
020 ← 010	0 – 62	0 – 21	446.390	622.055	3727	0.024	0.005	0.004	5.77×10^{-19}	5.94×10^{-19}	-38.7%	34.0%	2.6%	16.1%
030 ← 020	1 – 46	0 – 17	463.097	598.267	1532	0.028	0.015	0.007	5.59×10^{-20}	5.72×10^{-20}	-29.5%	26.0%	2.0%	12.6%
100 ← 000	0 – 88	0 – 32	1030.973	1273.175	8291	0.052	0.003	0.004	3.32×10^{-18}	3.63×10^{-18}	-42.3%	34.6%	6.4%	11.2%
110 ← 010	0 – 45	0 – 22	1047.859	1243.820	4043	0.021	0.005	0.004	2.51×10^{-19}	2.74×10^{-19}	-99.3%	27.4%	6.0%	18.1%
001 ← 000	0 – 90	0 – 32	1294.334	1409.983	5686	0.041	0.006	0.004	2.57×10^{-17}	2.79×10^{-17}	-47.5%	14.1%	1.2%	7.6%
011 ← 010	0 – 71	0 – 25	1302.056	1397.007	3948	0.034	0.014	0.006	2.02×10^{-18}	2.22×10^{-18}	-15.1%	36.0%	5.4%	6.7%
101 ← 000	0 – 82	0 – 24	2433.192	2533.195	4034	0.110	0.023	0.011	5.39×10^{-19}	5.34×10^{-19}	-41.7%	27.2%	-5.3%	5.2%
111 ← 010	0 – 61	0 – 21	2441.079	2521.117	2733	0.145	0.011	0.013	4.24×10^{-20}	4.25×10^{-20}	-30.5%	4.2%	-2.5%	4.1%
002 ← 000	0 – 76	⋆0 – 24	2599.080	2775.076	4327	0.033	0.011	0.006	3.77×10^{-21}	3.51×10^{-21}	-97.9%	57.9%	-10.8%	13.4%
003 ← 000	0 – 77	0 – 25	3985.185	4092.948	3655	0.122	0.031	0.030	1.55×10^{-21}	1.33×10^{-21}	-94.3%	21.3%	-31.3%	21.1%

⋆K_a = 11 excluded.

Figures 3.8 and 3.9 show the simulated cross sections for the $1000 < \nu < 1500$ cm^{-1} spectral region at 573.15 K (300 C) and 773.15 K (500 C), respectively, convolved with a Gaussian line shape function with HWHM = 0.25 cm^{-1}. These are compared with experimental cross sections obtained by Fateev *et al.* measured at a resolution of 0.5 cm^{-1}. The simulations are calculated using a cross section code, 'ExoCross', developed by S. Yurchenko designed to work with the ExoMol line list format [175], based on the principles outlined in Ref. [188].

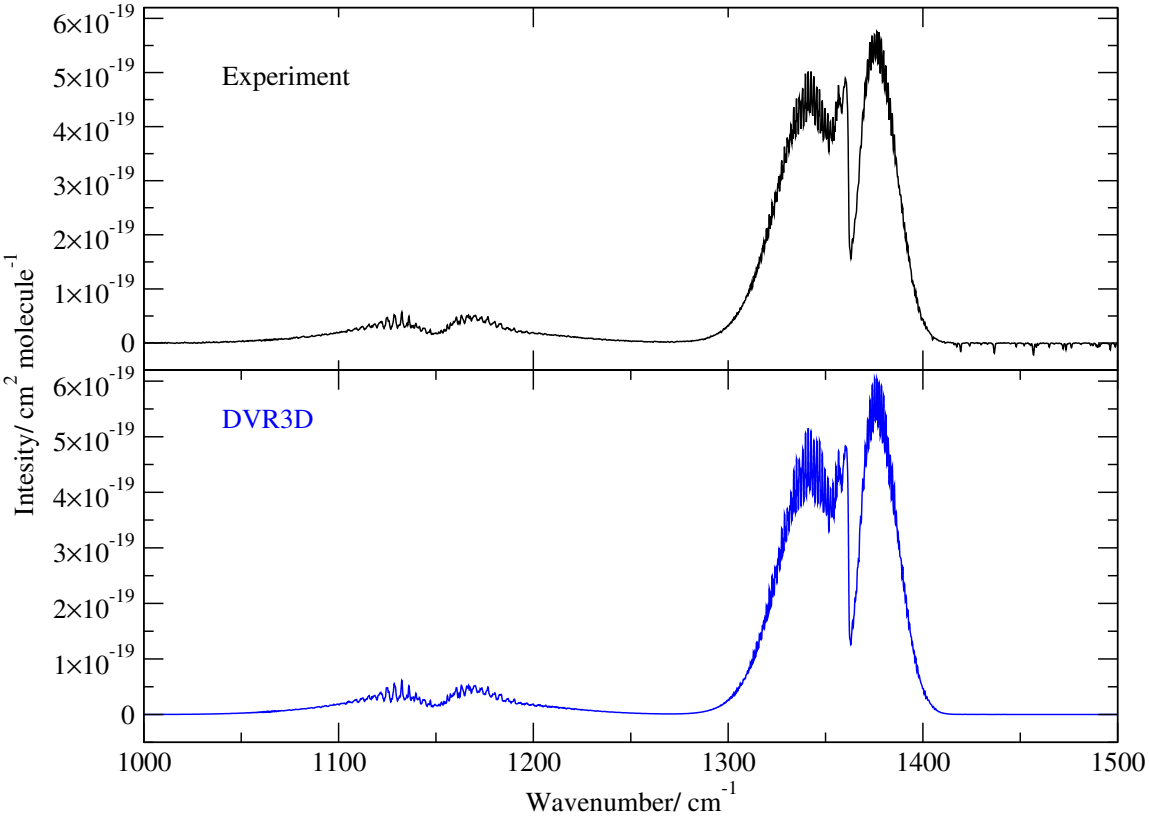

Figure 3.8: Comparison of cross sections obtained at T = 573.15 K for $^{32}S^{16}O_2$ via experimental observation by Fateev *et al.* [187] (above) and from the hot line list (below).

This spectral region contains both the ν_1 and ν_3 bands, and the intensity features are qualitatively well represented by the simulated cross sections. For 573.15 K (300 C) the integrated intensity across the $1000 < \nu < 1500$ cm^{-1} spectral region is calculated as 3.43×10^{-17} cm^2 molecule^{-1}, which is about 2% less than that for the experimental value, measured as 3.50×10^{-17} cm^2 molecule^{-1}.

For 773.15 K (500 C) the integrated intensity across the same spectral region is calculated as 3.41×10^{-17} cm^2 molecule^{-1}, which is roughly 6% less than that for the experimental value, 3.62×10^{-17} cm^2 molecule^{-1}. This may be attributed to a

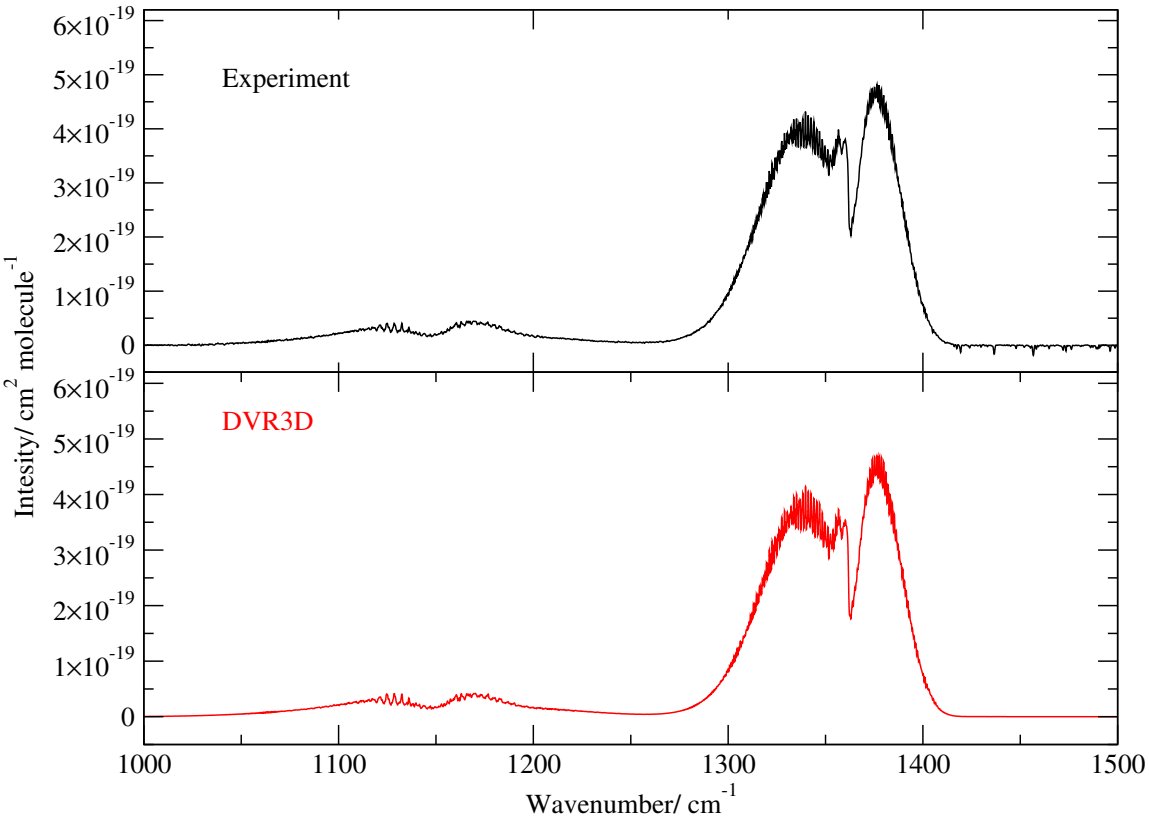

Figure 3.9: Comparison of cross sections obtained at T = 773.15 K for $^{32}S^{16}O_2$ via experimental observation by Fateev *et al.* [187] (above) and from the hot line list (below).

small discrepancy observed in the P-branch of the ν_3 band which is not obvious from Figure 3.9; the intensity here is slightly lower for the computed cross sections. Since this disagreement affects a specific region of the spectrum, it is unlikely wholly due to an error in the partition sum. The quality of the DMS may also be a contributing factor, in conjunction with the states involved in these transitions. Another source may be from the generation of the cross sections themselves; the line shape function used in constructing the theoretical cross sections is Gaussian, and therefore only considers thermal (Doppler) broadening, as opposed to a combination of thermal and pressure broadening (Voigt line shape). It is possible that neglecting the pressure broadening contribution in the line shape convolution is the source of this disagreement, as it is expected to have significance at elevated temperatures. Regardless of this discrepancy, using a Voigt profile would considerably improve the overall quality of computed cross sections. This is discussed further in Chapter 7.

3.4.4 Cross Sections

Figures 3.10 - 3.13 display temperature-dependent calculated cross sections for the rotational and fundamental bands of $^{32}S^{16}O_2$. All simulations are produced using the hot line list convolved with a Gaussian line shape function with HWHM = 0.5 cm^{-1}.

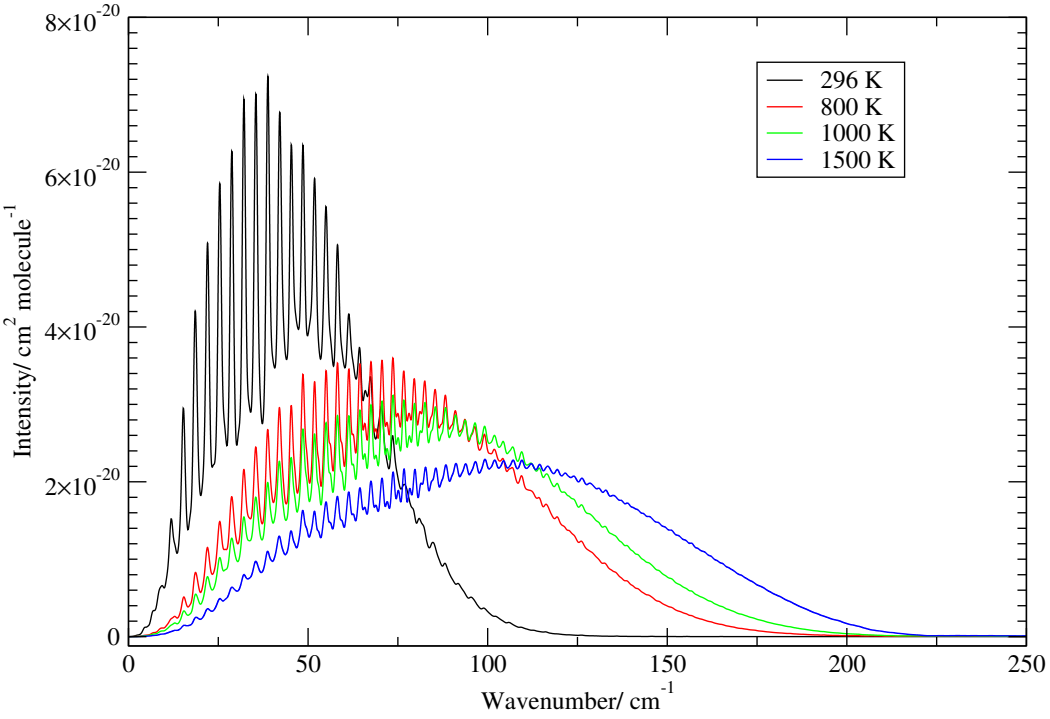

Figure 3.10: Temperature-dependent cross sections for the rotational band of $^{32}S^{16}O_2$.

Figure 3.14 shows an overview plot of the spectrum for $0 < \nu < 8000$ cm^{-1}, highlighting the temperature-dependence of the cross section intensities. Again, this simulation is produced using the hot line list convolved with a Gaussian line shape function with HWHM = 0.5 cm^{-1}.

3.5 Discussion

Application of the DVR3D method has been carried out in order to produce a hot line list for $^{32}S^{16}O_2$. The process has consisted of several adaptations to the pre-existing code in order to facilitate the high demands of this computation. The line list is based upon an *ab initio* PES and DMS computed by Huang *et al.* [35], which have both been shown to be correctly implemented into the DVR3D program by the comparison of rovibrational energies and transition intensities with existing data. The Ames-296K data is reproduced very well, and comparisons with HITRAN data for all $J = 0 - 99$

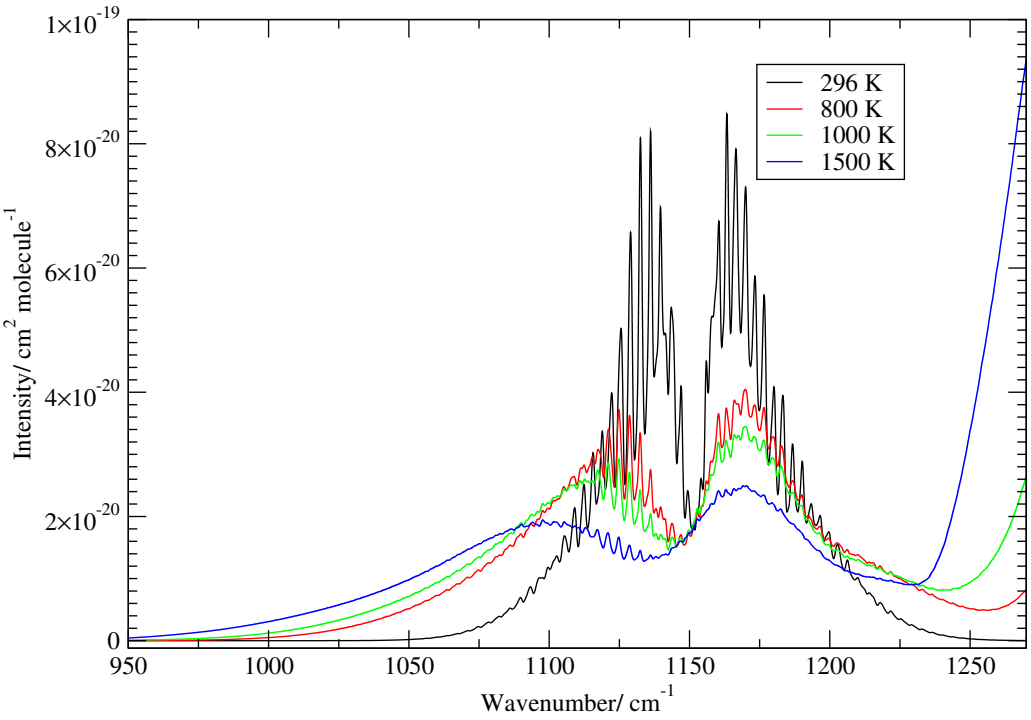

Figure 3.11: Temperature-dependent cross sections for the ν_1 band of $^{32}S^{16}O_2$. The contribution to the intensity beyond 1225 cm^{-1} is due to the ν_3 band.

rovibrational energy levels produces an overall root-mean-square error of 0.016 cm^{-1}. The majority of HITRAN intensity comparisons agree to within 80% - 90%, and provide similar correlation to comparisons drawn by Huang *et al.* Preliminary cross section data has been simulated and compared and has been shown to have good qualitative agreement with available experimental data, and may facilitate in the future analysis of hot spectra.

The main focus of this project has been in the production of fully converged wavefunctions suitable for the simulation of energies and transitions at high temperatures, which has been achieved. Discrepancies discussed in Ref. [35] are not necessarily alleviated by the convergence of higher energies, or by the application of a different computational method, and are inherent to both the PES and DMS. Huang *et al.* guarantee high-accuracy (0.01 – 0.03 cm^{-1}) for simulations below 5500 cm^{-1}, and therefore the same precautions should be followed here. Computed rovibrational energies above 10 000 cm^{-1}, though converged, can not be guaranteed to spectroscopically accurate, and are entirely dependent on the quality of the PES. This may have very minor repercussions on the convergence of the partition function at the high end of the temperature range.

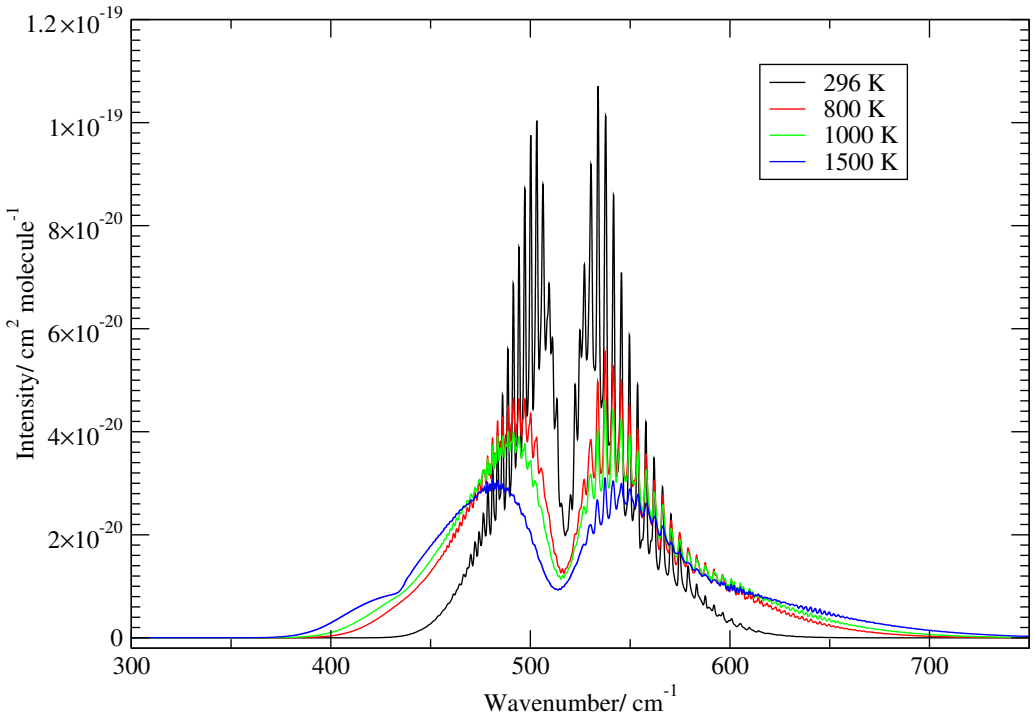

Figure 3.12: Temperature-dependent cross sections for the ν_2 band of $^{32}S^{16}O_2$.

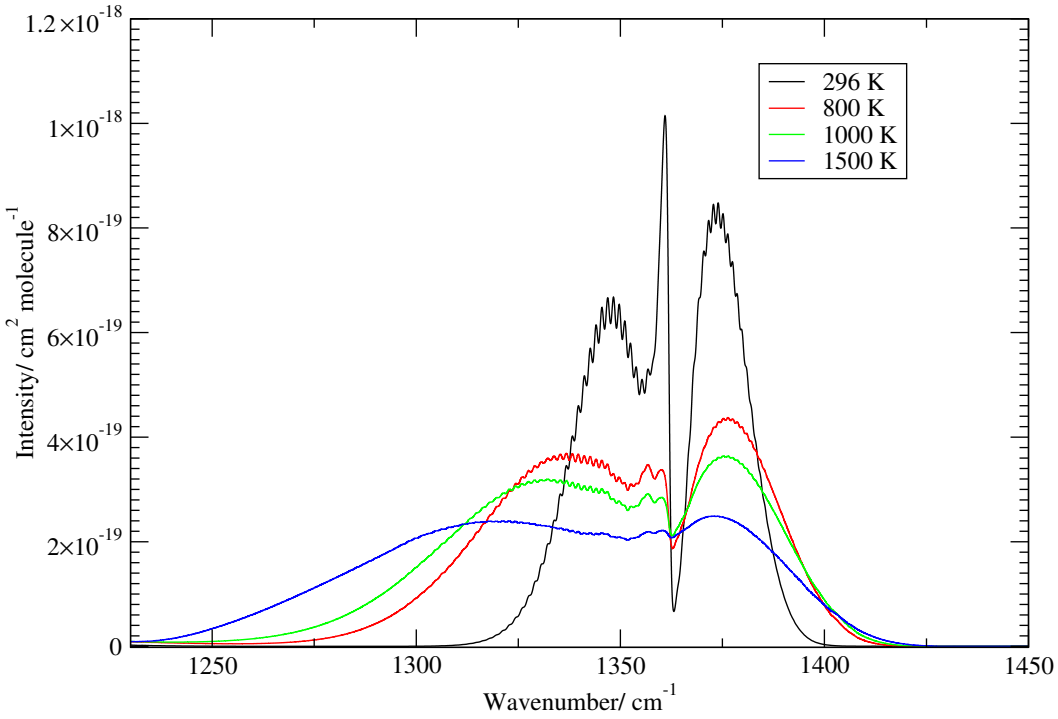

Figure 3.13: Temperature-dependent cross sections for the ν_3 band of $^{32}S^{16}O_2$.

Nevertheless, the scalability of the DVR3D suite has been greatly improved and is now well suited for future computations in conjunction with higher accuracy PESs and DMSs, which require the construction and storage of wavefunctions based on larger

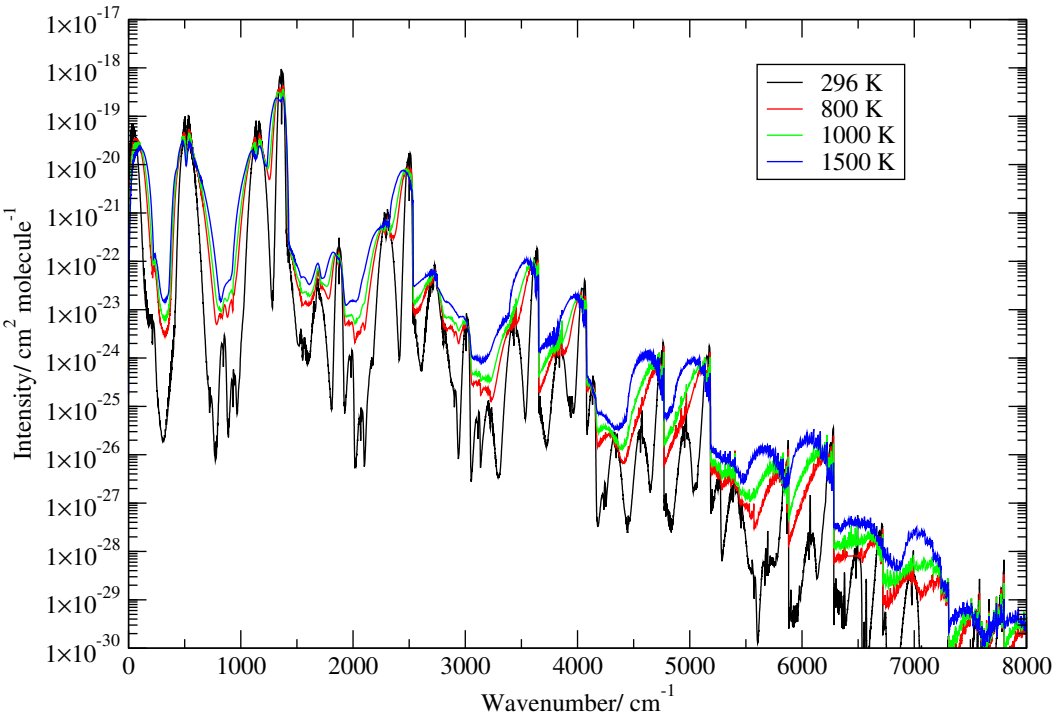

Figure 3.14: Temperature-dependent cross sections for the entire $0 < \nu < 8000$ cm^{-1} region of $^{32}S^{16}O_2$.

DVR grids. The final line lists consists of 1.3 billion transitions within $0 < \nu < 8000$ cm^{-1}, between 3 255 954 rovibrational energy levels up to $J = 165$. This line list is suitable for the simulation of rovibrational spectra for $T \leq 1500$ K, but may be used approximately for temperatures above this. The full line list can be downloaded from the ExoMol website (http://www.exomol.com).

Chapter 4

An Initial Room-Temperature Line List for $^{32}S^{16}O_3$

In this chapter the first stage of a theoretical study on spectroscopic data for SO_3 is presented. This is the first time a large, comprehensive theoretical investigation has been performed regarding this molecule. There have been a handful of previous studies performed for SO_3, both experimental and theoretical, but neither have attempted to cover both line position and transition strength data to the desired level. The aim of this project is to compute accurate line positions and Einstein coefficients complete within the chosen spectral window.

Since spectroscopic data for SO_3 is relatively sparse within the literature, with studies limited to room temperature analysis, it is difficult to test the reliability of line list data computed for higher temperatures. Both the potential and dipole moment surfaces for the current study are computed *ab initio*, with a view to refinement based on experimental data. For this reason the project is divided into two stages.

First, a room temperature line list for SO_3 is computed using the purely *ab initio* PES and DMS. This is compared to the available experimental data to determine any discrepancies observed in line positions; it is expected that the purely *ab initio* PES would not immediately produce accurate energy levels without an empirical refinement. This initial line list is usually computed at a reduced basis set size adequate for the coverage of data at lower temperatures, i.e. until partition function convergence is observed, and therefore is less computationally demanding than a high temperature line list.

The second stage of the full high-temperature line list computation involves the

empirical refinement of the PES, before the final calculation using an increased basis set and spectral range. This is the subject of discussion in Chapter 6.

Presented here are details of the first stage of the process, in computing the purely *ab initio* room-temperature line list, referred to as UYT (Underwood, Yurchenko, Tennyson). Discussed are the potential and dipole moment surfaces, convergence tests and variational calculations, along with comparisons between the final computed data and those available in the literature. All nuclear motion calculations were performed with the program TROVE [131] which had previously been adapted to work in the $\mathscr{D}_{3h}$(M) molecular permutation-inversion group for NH_3 [189], which belongs to the same point group as SO_3. The final line list consists of transitions on the region 0 - 4000 cm^{-1} with rotational states up to $J = 85$, resulting in 174 674 257 transitions.

4.1 Calculation using TROVE

The TROVE procedure [131] consists of a number of steps in order to calculate a line list, much like the DVR3D suite. The main difference in the case of TROVE is that a single program module is used in all steps, which reads in various different inputs which specify which calculation to perform, along with relevant input parameters. Each step produces a particular set of checkpoint files required for use in the following step(s). The general theoretical background describing the calculation procedure is outlined in Chapter 2. A summary of the computational steps are listed below:

Step 1 In the first step the expansion of the Hamiltonian is performed in terms of the chosen coordinates. This requires expanding the kinetic and potential energy operators, each to a particular order defined by the user input. The electric dipole moment function is also expanded in terms of the chosen coordinates. These expansions are stored as external checkpoint files to be read by later processes.

The primative 1D basis functions discussed in in Chapter 2 are also generated in this step. Only functions associated up to a certain number of vibrational quanta are considered, based on a truncation scheme with a limit specified in the user input. This truncation parameter is discussed further below in subsection 4.1.2. The 1D basis functions describe the internal motions of the molecule; in the case of SO_3 there are six internal vibrational degrees of freedom. The primitive functions are combined to produced a vibrational basis for subsequent

rovibrational calculations, the process is described in Chapter 2. The vibrational matrix elements are stored to disk, and the vibrational hamiltonian is diagonalised in this 6D product basis.

Step 2 The eigenfunctions computed in the previous step as solutions to the vibrational ($J = 0$) problem are now used to produce a compact basis set for all $J > 0$ calculations. This greatly facilitates the construction of $J > 0$ Hamiltonian matrices as their vibrational elements are now diagonal. The new rovibrational basis set is formed as a product between the $J = 0$ vibrational eigenfunctions and symmetrised rigid rotor wavefunctions. This stage of computation stores the vibrational and electric dipole moment matrix elements in terms of the new rovibrational basis set, to be used in further steps.

Step 3 Now eigenfunctions for $J > 0$ are computed. The Hamiltonian is built by reading the vibrational matrix elements in the $J = 0$ representation generated in the previous step and combining them with rigid rotor functions. As such the Hamiltonian matrix is block diagonal in J. This allows for the Hamiltonian to be built and diagonalised separately for individual J blocks. For each J an individual process computes the appropriate matrix. Since the underlying basis functions are symmetry labelled according to the Molecular Point Group (see Chapter 2), eigenfunctions of the rovibrational matrix are only required for states with physically allowed symmetry species. The required symmetries can be requested as user input; in the case of SO_3 the physically allowed symmetry species are A_1' and A_1'' in the $\mathscr{D}_{3h}$(M) molecular permutation-inversion group. The final computed eigenfunctions are saved to disk, stored in files grouping by J and the overall symmetry Γ_{Total}. Only wavefunctions with energies up to a user defined limit are saved.

Step 4 Once all necessary wavefunction files have been generated for J' and J'', TROVE can be used to generate absorption intensities, line strengths and Einstein coefficients for all transitions satisfying $J' \leftrightarrow J''$ and associated selection rules ($A_1' \leftrightarrow A_1''$), along with user-specified cut-offs for E'_{max} and E''_{max}. Lower thresholds for intensities/line strengths can also be specified, as well as the spectral window of interest, and the temperature and partition function can also be

entered, in order to generate the temperature-dependent line intensities. This is essentially the final TROVE step. All transitions of interest are computed with associated Einstein coefficients. Subsequent programs can use this data to generate a temperature-dependent spectrum.

4.1.1 The Potential Energy and Dipole Moment Surfaces

In order to produce rovibrational wavefunctions and energies using TROVE it is necessary to supply the program with an external potential surface. Provided the TROVE code has been adapted to work with the specific functional form of the PES the coefficients of this external function are given as input parameters to the first stage of the TROVE computation - the vibrational problem.

At the start of the project the only known fully *ab initio* work in the literature regarding a PES for SO_3 was by Martin [123]. Previous studies by Dorney *et al.* [121] provided force constants, which Martin expanded upon by computing an *ab initio* quartic force field developed using coupled cluster methods. The quadratic, cubic and quartic force constants were computed in symmetry-adapted internal coordinates. The very first calculations performed in this project, which involved various convergence tests (see next subsection) successfully implemented these constants. However a number of computed vibrational bands using this force field could not be reconciled with their experimental counterparts, an example being the ν_2 vibration (see next section). The production of a brand new *ab initio* PES was already under construction at this point, however, as improvements to electronic structure methods since the publication of the force field by Martin would improve upon the force constants. This new PES is the basis upon which rovibrational wavefunctions have been computed in the current project, and is discussed below.

4.1.1.1 The *ab initio* PES

The *ab initio* PES was constructed by Yurchenko [170] using the explicitly correlated F12 singles and doubles coupled cluster method including a perturbational estimate of connected triple excitations, CCSD(T)-F12b [190]. The basis sets used are correlation consistent and optimised for F12 calculations; the valence correlation-consistent functions aug-cc-pVTZ-F12 and aug-cc-pV(T+d)Z-F12 [191] are used for oxygen and sulphur, respectively. Molpro.2010 [192] was employed for all electronic structure

calculations. The masses for the sulphur and oxygen atoms used in all subsequent calculations are, respectively, 31.97207070 and 15.99491463 in amu.

A grid of 13 000 points was used with the stretching modes ranging from 1.1 to 2.0 Å, interbond angles α between 70 and 123°, and the inversion angle ρ between 0 and 50°. Grid points were specified based on an algorithm which produced a higher concentration of equidistant points close to the equilibrium position, but favoured less those that exceeded a certain energy value. Each point was processed with an MP2 calculation as a further filter, and those producing an MP2 result of 50 000cm^{-1} or higher were disregarded. This grid is sufficient to cover energies up to 40 000 cm^{-1} above the minimum. It is important to note that CCSD(T) is probably not reliable for the higher energy regions which approach dissociation. However, the nuclear motion calculations based upon this PES are largely sensitive to energies below 10 000cm^{-1}, and therefore the coupled cluster method used is assumed to be adequate.

This PES is represented by an analytical form given by the expansion [170, 189]

$$\begin{aligned} V(\xi_1,\xi_2,\xi_3,\xi_{4a},\xi_{4b};\sin\bar{\rho}) &= V_{\rm e}+V_0(\sin\bar{\rho})+\sum_j F_j(\sin\bar{\rho})\,\xi_j \\ &+ \sum_{j\le k} F_{jk}(\sin\bar{\rho})\,\xi_j\,\xi_k+\sum_{j\le k\le l} F_{jkl}(\sin\bar{\rho})\,\xi_j\,\xi_k\,\xi_l \\ &+ \sum_{j\le k\le l\le m} F_{jklm}(\sin\bar{\rho})\,\xi_j\,\xi_k\,\xi_l\,\xi_m \end{aligned} \tag{4.1}$$

in the coordinates ξ_k:

$$\xi_k = 1-\exp(-a(r_k-r_{\rm e})),\ k=1,2,3, \tag{4.2}$$

$$\xi_{4a} = \frac{1}{\sqrt{6}}(2\alpha_1-\alpha_2-\alpha_3), \tag{4.3}$$

$$\xi_{4b} = \frac{1}{\sqrt{2}}(\alpha_2-\alpha_3), \tag{4.4}$$

$$\sin\bar{\rho} = \frac{2}{\sqrt{3}}\sin[(\alpha_1+\alpha_2+\alpha_3)/6], \tag{4.5}$$

where

$$V_0(\sin\bar{\rho}) = \sum_{s=1} f_0^{(s)}(\sin\rho_e-\sin\bar{\rho})^s \tag{4.6}$$

and

$$F_{jk\ldots}(\sin\bar{\rho}) = \sum_{s=0} f_{jk\ldots}^{(s)}(\sin\rho_e-\sin\bar{\rho})^s. \tag{4.7}$$

This form has been used to represent the PESs of NH_3, PH_3, SbH_3, NH_3^+, and BiH_3 [193–196]. The potential parameters $f^{(s)}_{jk\ldots}$ were obtained through a weighted least squares fit to the *ab initio* points with a root-mean-squares (rms) error of 0.067 cm^{-1}. Weight factors were set using the expression suggested by Partridge and Schwenke [179]:

$$w_i = \frac{\tanh\left[-0.0005\ \text{cm} \times \left(V_i - 16000\ \text{cm}^{-1}\right)\right] + 1.002\,002\,002}{2.002\,002\,002\ \text{cm} \times V_i^{(w)}}, \tag{4.8}$$

where $V_i^{(w)} = \max(16000\ \text{cm}^{-1}, V_i)$, and V_i is the *ab initio* energy at the ith geometry (in cm^{-1}), measured relative to the equilibrium energy. These weight factors favor the energies below 16 000 cm^{-1}. The weighting function used was based on an empirical adjustment scheme used to ensure that the *ab initio* points were fitted to within their individual errors while obtaining accurate replication of experimental line positions. In conclusion, a total of 304 coefficients were determined to be readily input into the TROVE program.

The *ab initio* equilibrium geometry r_e as obtained from the fit is 1.42039 Å, compared to the experimentally derived value 1.41732 Å [99]. The experimental rotational energies of SO_3 [104] are best described by the latter value, and therefore is used in all rovibrational calculations in place of the *ab initio* value. Replacing the equilibrium structure may cause undesirably large changes to the vibrational energies, as, for example, the bending kinetic energy depends on r_e [197]. In order to minimise this effect the following procedure was employed.

The constructed PES can be considered as an expansion around $r_e^{(ai)}$, however the desired PES should be centred on $r_e^{(exp)}$. If the expansion is performed around $r_e^{(exp)}$ then by definition it would produce a completely different PES (PES-a) to the one calculated, whose structure would differ by a number of linear terms. In order to maintain the structure of the *ab initio* PES while ensuring an expansion around $r_e^{(exp)}$, the quantity $\Delta r_e = r_e^{(ai)} - r_e^{(exp)}$ is defined. A second value, $r_e^{(ref)}$ is defined as the point which is a distance Δr_e from $r_e^{(ai)}$, but in the opposite sense from $r_e^{(exp)}$. Expanding the PES around $r_e^{(ref)}$ defines yet another PES (PES-b), which also differs to the original PES by a number of linear terms. However substituting $r_e^{(ref)}$ with $r_e^{(exp)}$ has the effect of removing the linear terms in PES-a, therefore maintaining the structure of the original

PES while expanding around $r_e^{(exp)}$.

Table 4.1 compares rotational levels measured by Maki *et al.* [104] and those computed using the equilibrium-adjusted PES used in this work. The agreement is very good.

4.1.1.2 The *ab initio* DMS

The DMS was calculated using the same level of theory as the PES and on the same grid of 13 000 points. The *ab initio* values were then expressed analytically using a symmetrized molecular bond (SMB) representation described in Ref. [198]. The resulting dipole moment parameters obtained through a least squares fit reproduce the *ab initio* data with an rms error of 0.00013 D. In these fits the same factors defined by Eq. (4.8) were used to weight the geometries according to the corresponding energies. As with the PES, a total of 304 coefficients were determined to be readily inputted into the TROVE program. At the time this project was carried out, there were no experimental or *ab initio* dipole moment data in the literature that could be use to validate the *ab initio* DMS against. However, experience of working dipole moments for different systems [198–201] has shown that high quality *ab initio* intensities in most cases are competitive with experimental measurements. Nevertheless, an attempt has been made to assess the quality of the relative intensities calculated using the *ab initio* DMS, and this is discussed below.

4.1.2 Convergence Testing

As with the SO_2 calculations described in Chapter 3, in order to achieve results of high accuracy as well as minimising the requirement for computational resources, it is necessary to optimise the size of the nuclear motion Hamiltonian matrix. This involves preliminary truncation of the basis set, as well as limiting the order of both the kinetic and potential components of the Hamiltonian expansion [131]. Deciding upon these expansion orders involves a trade-off between accuracy and computational expense. These expansion orders were the first to be determined in the set of convergence tests.

A number of combinations for the expansion orders were tested with N_{kin} = 4, 6, 8, and with N_{pot} = 4, 6, 8, 10. It was found that convergence in the kinetic energy expansion occurred before that of the potential energy expansion, and ultimately setting

Table 4.1: Theoretical rotational term values for $^{32}S^{16}O_3$ ground vibrational state (cm^{-1}) compared with experiment [104].

J	K	Obs.	TROVE
2	0	2.0912	2.0916
3	3	2.6115	2.6119
4	3	5.3998	5.4006
4	0	6.9707	6.9718
5	3	8.8852	8.8864
6	6	8.3548	8.3559
6	3	13.0675	13.0694
7	6	13.2342	13.2360
7	3	17.9467	17.9493
8	6	18.8106	18.8132
8	3	23.5228	23.5263
8	0	25.0935	25.0972
9	9	17.2297	17.2319
9	6	25.0838	25.0874
9	3	29.7958	29.8002
10	10	17.2297	17.2319
10	9	24.2002	24.2035
10	6	32.0539	32.0584
10	3	36.7655	36.7709
10	0	38.3360	38.3417
20	18	89.8252	89.8372
20	15	107.0973	107.1122
20	12	121.2253	121.2425
20	9	132.2114	132.2304
20	6	140.0574	140.0777
20	3	144.7645	144.7856
20	0	146.3334	146.3548
80	78	1195.6589	1195.8085

$N_{kin} = 4$ and $N_{pot} = 8$ yielded adequately converged results at minimal computational cost; using a kinetic energy expansion order of 6 requires a more expensive calculation where convergence is already observed to within 0.001 cm^{-1} when expanding to fourth order. Table 4.2 shows computed fundamental bands using the quartic force field potential by Martin [123] using varying expansion order combinations for N_{kin} and N_{pot}; the converged values of the fundamental bands are not spectroscopically accurate, however this is not necessary to observe convergence. Table 4.3 shows the summation of the first 200, 400, 600, 800, 1000, 1200, and 1508 eigenvalues computed at different kinetic/potential expansion orders, in an effort to show the convergence.

After determining the expansion orders for the kinetic and potential energy com-

Table 4.2: Convergence of basis set for kinetic and potential expansions: vibrational band centres (cm^{-1}) for $^{32}S^{16}O_3$.

N_{kin}, N_{pot}	ν_1	ν_2	ν_3	ν_4
4, 4	1063.285404	428.341032	1386.453660	527.177358
4, 6	1063.360753	428.360347	1386.809831	527.349671
4, 8	1063.361192	428.360340	1386.810051	527.349228
6, 6	1063.360296	428.359962	1386.809558	527.349590
6, 8	1063.360736	428.359995	1386.809778	527.349146
8, 10	1063.360725	428.359950	1386.809774	527.349146

Table 4.3: Convergence of basis set for kinetic and potential expansions: summing of eigenvalues (cm^{-1}) for $^{32}S^{16}O_3$.

N_E	$N_{kin} = 6, N_{pot} = 8$	$N_{kin} = 8, N_{pot} = 10$
200	851761.14	851761.13
400	2109965.2	2109965.15
600	3538487.59	3538487.44
800	5111050.14	5111049.78
1000	6805390.58	6805390.04
1200	8616196.34	8616195.55
1508	11730399.33	11730398.2

ponents of the Hamiltonian, it is necessary to determine the size of the basis set used in the subsequent calculations. TROVE employs a polyad number truncation which controls the size of the basis set. For SO_3 the polyad number is given by

$$P = 2(n_1 + n_2 + n_3) + n_4 + n_5 + \frac{n_6}{2}, \tag{4.9}$$

where n_i are the quanta associated with 1D basis functions, ϕ_i, whose product gives the vibrational basis set [131]. Each of these basis functions is associated with an internal coordinate ξ_i, and only functions for which $P \leq P_{\max}$ are included in the primitive basis set. Initial tests were carried out to measure the degree of convergence using different values for $P_{\max}$, and the orders of the kinetic and potential energy expansions. In the present study, it is apparent that the convergence is more sensitive to $P_{\max}$, and convergence is obtained to within 0.1 cm^{-1} when $P_{\max}$ is 12 (see table 4.4); therefore a basis set based on this value is used.

As well as using the polyad number to truncate the size of the basis, a further truncation technique is employed by specifying an upper limit for the eigenvalue calculations, i.e. construct the basis set such that it provides energy values up to a limit

Table 4.4: Convergence of basis set viewed for some vibrational band centres (cm^{-1}) for $^{32}S^{16}O_3$.

	Obs. [104]	P_{max}= 10	P_{max}= 12	P_{max}= 14
ν_1	1064.92	1065.83	1065.75	1065.74
ν_2	497.57	498.48	498.48	498.48
ν_3	1391.52	1387.63	1387.45	1387.43
ν_4	530.09	528.61	528.59	528.58
$2\nu_3(l_3$=0)	2766.40	2759.61	2759.12	2758.75
$2\nu_3(l_3$=2)	2777.87	2770.70	2770.29	2769.95
$2\nu_2$	995.02	995.43	995.35	995.35
$2\nu_4(l_4$=0)	1059.81	1057.10	1056.50	1056.44
$2\nu_4(l_4$=2)	1060.45	1057.86	1057.38	1057.33
$\nu_2 + \nu_4(l_4$=1)	1027.90	1027.58	1027.35	1027.33
$\nu_1 + \nu_4(l_4$=1)	1593.69	1593.82	1593.36	1593.30
$3\nu_4(l_4$=1)	1589.81	1587.64	1586.46	1586.30
$3\nu_4(l_4$=3)	1591.10	1587.61	1586.43	1586.27
$\nu_1 + \nu_2$	1560.60	1565.51	1565.33	1565.32
$\nu_2 + 2\nu_4(l_4$=0)	1557.88	1556.38	1555.59	1555.47
$\nu_2 + 2\nu_4(l_4$=2)	1558.52	1557.12	1556.45	1556.37
$2\nu_2 + \nu_4(l_4$=1)	1525.61	1524.81	1524.48	1524.46
$3\nu_2$	1492.35	1449.81	1490.76	1490.76
$\nu_2 + \nu_3(l_3$=1)	1884.57	1881.82	1881.53	1881.51
$3\nu_3(l_3$=1)	4136.39	4138.88	4126.78	4125.92

of E_{max}. This is based on an estimate given by summing eigenvalues of the 1D basis functions before they are considered for matrix element calculations; the active space is constructed using basis functions whose eigenvalues sum together to have $E \leq E_{max}$. For the present study the condition E_{max}/hc = 10 000 cm^{-1} is used.

These precautions are particularly important for the SO_3 molecule, as its larger mass (compared to, for example, XH_3 systems) gives rise to small rotational constants, which in turn requires calculations up to high J value to ensure adequate coverage of transitions for a given temperature. This means that use of any unnecessary basis functions will prove computationally costly.

As an intial test of the TROVE procedure the quartic force field given by Martin [123] was used to compute fundamental term values using a well-converged polyad truncation scheme of P_{max}= 16. Discrepancies were found between the results and the values published, particularly the value of the ν_2 fundamental term value. A substition of the symmetry-adapted force constant F_{22} associated with this vibration with a scaled value from a previously published force-field [121] resulted in a significant improve-

ment. Table 4.5 illustrates the differences between these values, and the final converged result using the PES by Yurchenko for the fundamental vibrations, and it can be seen that convergence is observed using to an improved degree using the PES calculated in the current project.

Table 4.5: A comparison of the fundamental term values (cm^{-1}) between this work and Martin [123] (computed with TROVE). The second 'Martin F_{22}' column refers to a calculation using a substituted F_{22} value (see text).

	Obs. [104]	Martin	Martin F_{22} [121]	This work
ν_1	1064.92	1063.36	1064.22	1065.75
ν_2	497.57	428.36	487.10	498.48
ν_3	1391.52	1386.81	1386.85	1387.45
ν_4	530.09	527.35	527.32	528.59

4.1.3 Computing the Line List

In the first step of the TROVE calculation where the vibrational basis is constructed, the 1D functions ϕ_i (i = 1, 2, 3) are Numerov-Cooley derived functions for the three stretching vibrations, in terms of linearised stretching coordinates $\Delta r = R_i - r_e$ (i = 1, 2, 3), where R_i is the length of one of the three S - O bonds. For the bending functions, harmonic-oscillator functions are used in terms of linearised bending coordinates $\Delta\alpha_i$ (i = 1, 2, 3) . The out-of-plane bending functions are also Numerov-Cooley functions based on the inversion angle ρ [131]. The kinetic energy operator is expanded in terms of the ξ_i (i = 1 - 6) linearised coordinates for all internal motions. In the potential energy expansion, Morse-like functions are used as expansion functions for the three stretches, whereas the expansion variables for the bending motions are the linearised coordinates ξ_i (i = 4 - 6). More details can be found in Chapter 2 and in Ref. [131].

In addition to the basis set minimisation described in the previous subsection, it is possible to reduce the size of the Hamiltonian further by making use of group theory. SO_3 has $\mathscr{D}_{3h}$(M) molecular group symmetry and the spin-0 Bosons which make up the constituent atoms of this molecule allow the molecular rovibrational wavefunctions to exhibit the symmetry of only two of the six irreducible representations of this group in order to satisfy the Pauli Principle; namely the A_1' and A_1'' representations. This reduces both the number of Hamiltonian matrices we need to consider and, since E symmetry Hamiltonian matrices are larger, their size (see Chapter 2).

Wavefunctions were calculated for all rovibrational states up to J = 85 with A_1' and A_1'' symmetry. All wavefunctions corresponding to rovibrational energies below a threshold energy of 10,000 $cm^{-1}$are saved. In computing dipole transition moments between rovibrational states with E' and E'' for the upper and lower energies, respectively, thresholds are introduced for these energies to control the number of transitions computed, as well as defining limits for the spectral range of interest. In the room-temperature calculation all transitions satisfying $E'_{max} \leq 7000\,cm^{-1}$, $E''_{max} \leq 3000\,cm^{-1}$ and $0 < \nu \leq 4000\,cm^{-1}$ are computed. This clearly does not make use of all the computed energies, however a converged value in the partition function can be ensured by including these higher energy states in the partition sum (see next section). The transition dipoles, line strengths S, Einstein A coefficients and intensities I for allowed transitions are computed. For $^{32}S^{16}O_3$, the rigorous selection rules determining allowed electric dipole transitions are $\Delta J = J' - J'' = 0, \pm 1$ $(J'' + J' \geq 1)$, and symmetry selection rules $A_1' \leftrightarrow A_1''$.

All calculations were carried out on local multi-CPU computing machines in the Unversity College London Department of Physics and Astronomy, each equipped with Intel Xeon processors of either 4 or 8 cores. Memory resources ranged from 16 to 96 Gb between machines. The compiled TROVE code was optimised for OpenMP parallelisation which allowed for matrix diagonalisations and dipole transition moments to be calculated across multiple CPUs. The total time of the line list computation was roughly 13,100 CPU hours. Due to the limitations imposed by the usage of machines by multiple users, this amounted to roughly 6-8 months of real time. A total of 176,674,257 transitions are obtained for room temperature (296/298.15 K) from the calculations. Energy levels are labelled by a unique set of quantum numbers, including total symmetry in $\mathscr{D}_{3h}$(M), the vibrational symmetry in $\mathscr{D}_{3h}$(M), the rotational quantum numbers J and K and the vibrational local mode quantum numbers.

4.2 The Partition Function

As well as the spectral region of interest, the amount of rovibrational energies required also depends on the desired temperature at which to simulate a spectrum. The quality of a computed spectrum will become sensitive to $E_{\max}$ as the temperature increases; it is necessary to ensure that all energy states that are significantly populated for the

given temperature are calculated. This can be checked using the temperature-dependent partition function:

$$Q = \sum_i g_i \exp(-E_i/kT), \tag{4.10}$$

where g_i is the total degeneracy of the ro-vibrational state i with energy E_i, with the sum running over all energies at the absolute temperature T, and k is Boltzmann's constant. The total degeneracy is given by $(2J+1)$ times the nuclear spin degeneracy, which for the present case of $^{32}S^{16}O_3$ is simply 1 for both the A_1' and A_1'' symmetries, given that the nuclear spin of ^{16}O and ^{32}S are zero.

For a given temperature, it is possible to determine the contribution of various states to the value of Q and then check that Q converges to a specific value as E_i tends to infinity; as T increases a greater coverage of higher-lying energy states is required. For T = 298.15 K calculations show that Q converges to better than 1 % at $J = 85$, with a value of $Q = 7908.266$. Therefore calculations spanning all J's up to 85 should be sufficient for simulating spectra at this temperature. Figure 4.1 shows the value of Q as a function of all energy levels having J quantum number up to a maximum value, J_{max}, at an absolute temperature of T = 298.15 K. As energies are included in the summation for increasing values of J, associated energy levels contribute less and less to the value of Q, until it converges to a limit.

4.3 Intensity Simulations

In order to determine the validity of the calculated line list it is necessary to compare simulated spectra with spectra obtained through experimental procedures. In the present case it is expected that the purely *ab initio* PES will deliver inaccurate line positions, and the need to empirically refine the PES must be dealt with. However it is still possible to test the accuracy of the *ab initio* DMS through comparisons of transition moments/absolute intensities, provided that both *ab initio* and experimental transitions can be adequately defined by upper and lower state quantum numbers. Since there is a lack of data in the literature regarding dipole moments, the best mechanism for comparison lies in computing absolute intensities at temperatures for which absolute or relative intensities are available. This is the approach taken here.

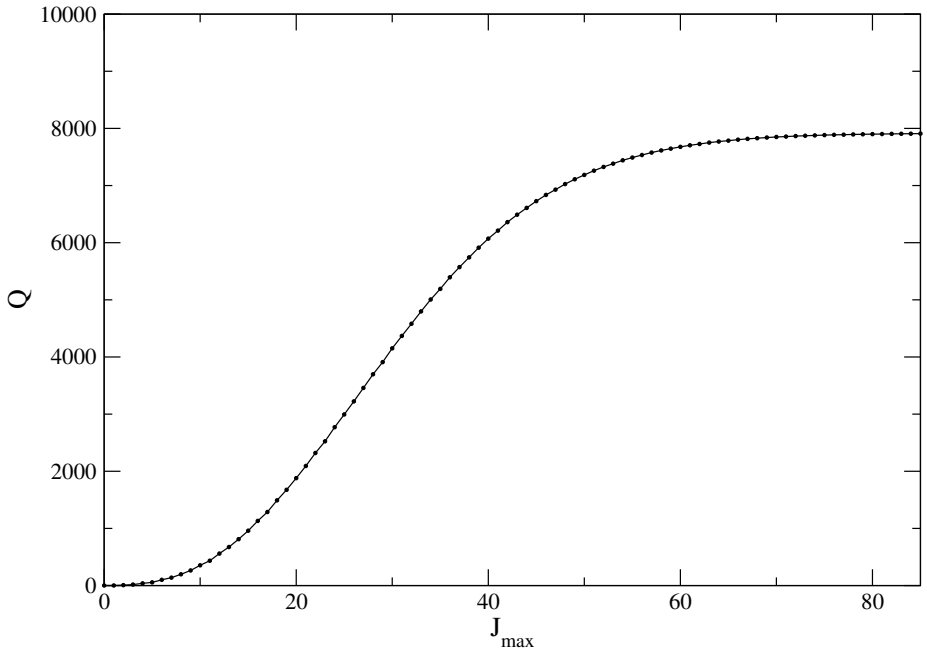

Figure 4.1: Convergence of partition function at 298.15 K for different values of $J_{\max}$.

The intensity of a transition between two states is given by the formula [202]

$$I(f \leftarrow i) = \frac{8\pi^3 N_A \nu_{if}}{(4\pi\varepsilon_0)3hc} \frac{e^{-E_i/kT}}{Q(T)} \times \left[1 - \exp\left(\frac{-hc\nu_{if}}{kT}\right)\right] S(f \leftarrow i), \tag{4.11}$$

where $I(f \leftarrow i)$ is the transition intensity for a transition from state i with energy E_i to state f with energy E_f, with $hc\nu_{if} = E_f - E_i$. Q is the partition function. $S(f \leftarrow i)$ is the line strength, which is defined by the following integration [202]:

$$S(f \leftarrow i) = g_{\rm ns} \sum_{m_f, m_i} \sum_{A=X,Y,Z} |\langle \Phi_{\rm rv}^{(f)} | \mu_A | \Phi_{\rm rv}^{(i)} \rangle|^2 \tag{4.12}$$

for a transition between initial state i with rovibrational wavefunction $\Phi_{\rm rv}^{(i)}$ and final state f with wavefunction $\Phi_{\rm rv}^{(f)}$. Here, $g_{\rm ns}$ is the nuclear spin statistical weight factor, and μ_A is the electronically averaged component of the molecular dipole moment along the space-fixed axis $A = X, Y, Z$. The sum also runs over the quantum numbers m_i and m_f which are projections of the total angular momentum J (in units of $\hbar$) on the laboratory fixed Z axis, for the initial and final states, respectively.

Maki *et al.* [99–105] reported extensive high-resolution studies of a number of fundamental, combination and overtone bands of $^{32}S^{16}O_3$. Their principle aim was to obtain accurate wavenumber measurements, but relative intensities were also measured.

In the present work these data are converted into absolute intensities by normalizing to the theoretical intensities obtained with TROVE at T = 298.15 K as described below.

The experimental measurements available cover three spectral regions: 405–708 cm^{-1} (focusing on ν_2, ν_4, $2\nu_2$-ν_2, ν_2+ν_4-ν_2, ν_2+ν_4-ν_4, ν_1-ν_4, $2\nu_4^{(l_4=0)}$-ν_4 and $2\nu_4^{(l_4=2)} - \nu_4$), 1200–1680 cm^{-1}(ν_3), and 2500-3280 cm^{-1} ($2\nu_3^{(l_3=2)}$). Each measurement was made at different values of pressure. For the 405–708 cm^{-1} window, the measurements were performed at 0.409 and 2.04 Torr, and for the ν_3 measurements between the 1200–1680 cm^{-1} window 0.16 Torr and 0.7 Torr was used. 560 lines were measured at 0.7 Torr, however 439 of these had relative intensity values which were negative. Therefore these higher pressure measurement were not used at all in the present work.

Table 4.6 shows the comparison between the numbers of lines identified in each measurement to the numbers of lines computed using TROVE. The latter numbers are the subject of the following selection criteria: $J \leq 85$, intensity cut-off, $I(f \leftarrow i) > 10^{-34}$ cm/molecule, and the wavenumber window, 0–4000 cm^{-1}. Experimental lines with negative relative intensities were left out of the analysis.

To normalize the experimental intensities, the experimental relative data from each spectral window and each different pressure were scaled to match the theoretical values computed at T = 298.15 K. The scaling factors obtained through a minimization procedure using all selected experimental lines with non-zero intensity (see Table 4.7) are 6.571×10^{-21} cm/molecule (405–708 cm^{-1}, 0.409 Torr), 1.838×10^{-21} cm/molecule (405–708 cm^{-1}, 2.04 Torr), 4.823×10^{-20} cm/molecule (1200–1680 cm^{-1}, 0.16 Torr), and 1.328×10^{-21} cm/molecule (2500-3280 cm^{-1}, 4.99 Torr). With these factors the 'experimental' intensities match the theoretical values reasonably well, for example for 405–708 cm^{-1} the agreement is within about 7.02×10^{-22} cm/molecule and 5.05×10^{-22} cm/molecule, at 0.409 and 2.04 Torr, respectively.

Having the absolute intensities derived, band intensities were estimated for each experimental band as the sum of individual line intensities. In Table 4.7 these 'experimental' band intensities S^{exp} (cm/molecule) are compared to the theoretical values obtained by summing intensities (a) from all TROVE lines from a given window and (b) only from lines with experimental counterparts present. This was done separately for each spectral range, and each measurement pressure therein. In Table 4.7 these quan-

tities are referenced to as $S_{\text{tot}}^{\text{calc}}$ and $S_{\text{red}}^{\text{calc}}$ for the 'total' and 'reduced' band intensities, respectively, and compared to S^{exp}.

The ratio $S_{\text{red}}^{\text{calc}}$ to S^{exp} also shown in Table 4.7 demonstrates the good quality of the scaling procedure employed as well as that of the computed *ab initio* dipole moment. For example at 0.409 Torr, the differences between 'experimental' and theoretical band intensities are within about 20% for all bands from the 405–708 cm^{-1} region with the exception of $2\nu_4^{(l_4=0)}$ - ν_4 (see also discussion below). It should be stressed here that only one scaling factor for all eight bands from this window was applied at a given pressure. The difference between the two theoretical band intensities $S_{\text{tot}}^{\text{calc}}$ and $S_{\text{red}}^{\text{calc}}$ gives a measure of the missing experimental transition data. According to Table 4.7 the measurements for even stronger bands are missing more that 50% of the total intensity.

Table 4.7 also shows theoretical values of vibrational transition moments defined as

$$\bar{\mu} = \sqrt{\bar{\mu}_x^2 + \bar{\mu}_y^2 + \bar{\mu}_z^2}, \tag{4.13}$$

where

$$\bar{\mu}_\alpha = \langle \Psi_{\text{vib}}^{(i)} | \mu_\alpha | \Psi_{\text{vib}}^{(f)} \rangle, \tag{4.14}$$

and $\Psi_{\text{vib}}^{(i)}$ and $\Psi_{\text{vib}}^{(f)}$ are the vibrational eigenfunctions of the 'initial' and 'final' states, respectively, variationally computed using TROVE and μ_α is the component of the molecular dipole moment along the molecular-fixed axis $\alpha = x, y, z$.

Figure 4.2 presents an overview of the simulated spectrum (T = 298.15 K) with TROVE and experimental absorption spectra of SO_3 for the whole simulation range up to 4000 cm^{-1}. It reveals the gaps and limitations of the available experimental data. Figure 4.3 shows the 'forbidden' rotational band as a stick spectrum. The characteristics of this band are discussed later in Chapter 5. It should be noted that the microwave measurements of Ref. [106] do not report any intensities. In Figure 4.4 all eight bands from the 405–708 cm^{-1} region are combined into one graph for each pressure to illustrate the quality of the corresponding experimental data. This figure suggests that the 0.409 Torr data are generally more reliable. This is also reflected by the ratio values $S_{\text{red}}^{\text{calc}}/S^{\text{exp}}$ from Table 4.7, which are significantly closer to 1 at the lower pressure. For data obtained at two pressures, scaled intensity values obtained for the lower pressure are deemed to be more reliable. Finally, Figure 4.5 presents a detailed comparison

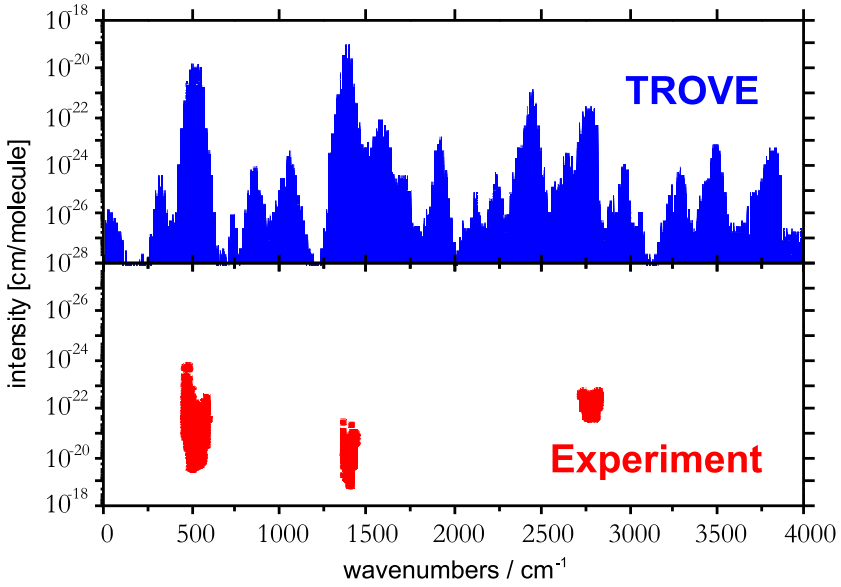

Figure 4.2: Overview of the simulated absorption (T = 298.15K) spectrum (TROVE) of SO_3 compared to experiment scaled to the theoretical intensities (see text).

for all the bands from the three spectral regions studied in this work in the form of stick diagrams. The intensities based on the *ab initio* DMS in TROVE are in very good qualitative agreement with the experiment.

Table 4.6: Comparison of calculated (TROVE) and experimental [104] band centers and numbers of line transitions.

Band	Obs.	Calc.	P_1	P_2	TROVE
ν_2 - ν_0	497.57	498.48	773	1265	5422
ν_4 - ν_0	530.09	528.59	996	2052	12195
ν_1 - ν_4	534.83	537.16	0	69	15147
$\nu_2 + \nu_4$ - ν_2	530.33	528.87	84	571	12477
$2\nu_2$ - ν_2	497.45	496.88	112	704	7171
$\nu_2 + \nu_4$ -ν_4	497.81	498.76	47	602	27182
$2\nu_4^{(l_4=2)} - \nu_4$	530.36	528.79	116	775	31096
$2\nu_4^{(l_4=0)} - \nu_4$	529.72	527.91	39	455	13718
ν_3 - ν_0	1391.52	1387.45	2014	–	14441
$2\nu_3^{(l_3=2)} - \nu_0$	2777.87	2770.29	1527	–	18115
ν_0 - ν_0	–	–	25	–	3439

For the the 405 - 708 cm^{-1} region, measurements were made at two pressures; P_1 = 0.409 Torr, and P_2 = 2.04 Torr. For the ν_3 and $2\nu_3^{(l_3=2)}$ bands, measurements were taken at 0.16 Torr and 4.99, Torr respectively (Maki *et al.* [104]). Pressure values are not recorded for microwave measurements [106].

Table 4.7: Vibrational band intensities S^{exp}, $S^{\text{calc}}_{\text{tot}}$, $S^{\text{calc}}_{\text{red}}$ in cm/molecule$\times 10^{-18}$, and calculated transition moments $\bar{\mu}_{\text{if}}$ in Debye. P_1 and P_2 refer to the different pressure measurements within the same wavenumber region (see Table 4.6). $S^{\text{calc}}_{\text{red}}/S^{\text{exp}}$ is the ratio of the theoretical reduced and total band intensities (see text). $S^{\text{calc}}_{\text{tot}}$ is the theoretical band intensity computed by summing all TROVE lines. S^{exp} is the experimental band intensity obtained from a summation over all experimental values after scaling factors applied (see text). $S^{\text{calc}}_{\text{red}}$ is the theoretical band intensity computed using only lines for which experimental counterparts exist. N_{red} is the number of matched lines.

Band	S^{exp}		$S^{\text{calc}}_{\text{tot}}$	$S^{\text{calc}}_{\text{red}}$		$S^{\text{calc}}_{\text{red}}/S^{\text{exp}}$		N_{red}		$\bar{\mu}_{\text{if}}/D$
	P_1	P_2		P_1	P_2	P_1	P_2	P_1	P_2	
ν_2	2.987	1.537	3.705	2.915	1.559	0.98	1.01	773	1265	0.158
ν_4	4.258	3.112	5.949	4.310	3.149	1.01	1.01	995	2052	0.200
$2\nu_2$ - ν_2	0.116	0.470	0.661	0.101	0.411	0.88	0.88	112	704	0.221
$\nu_2+\nu_4$ - ν_2	0.062	0.322	0.528	0.052	0.251	0.84	0.78	84	571	0.199
$\nu_2+\nu_4$ - ν_4	0.026	0.260	0.581	0.022	0.215	0.84	0.83	47	602	0.223
$2\nu_4(l_4=2)$ - ν_4	0.112	0.589	0.873	0.102	0.485	0.91	0.82	116	769	0.283
$2\nu_4(l_4=0)$ - ν_4	0.026	0.222	0.405	0.015	0.179	0.57	0.81	38	454	0.196
ν_1 - ν_4	–	0.009	0.101	–	0.003	–	0.29	–	69	0.039
ν_3	39.490	–	44.440	39.490	–	–	–	2014	–	0.321
$2\nu_3$	0.093	–	0.119	0.093	–	–	–	1527	–	0.012

4.4 Discussion

The shifted theoretical bands in Figures 4.4 and 4.5 indicate that the computed *ab initio* PES of SO_3 requires improvement (see also band centers in Table 4.6). The spread of deviations is apparent; for example the theoretical ν_2 frequencies have an rms deviation of 0.77 cm^{-1} when compared with the experimental data, which is relatively small compared to the deviation for ν_3 of 4.07 cm^{-1}. Since the PES was computed completely *ab initio* and at a modest level of theory, this is to be expected. An empirical fit of the PES is necessary for high temperature calculations to achieve spectroscopically accurate energies, and this is discussed in Chapter 6.

Table 4.7 outlines the quality of the intensity scaling procedure. For the 0.409 Torr and 2.04 Torr measurements the comparisons mostly agree to within 20%, with the exception of the $2\nu_4^{(l_4=0)}$ - ν_4 band measured at 0.409 Torr which shows nearly a 50% difference, and the ν_1 - ν_4 band measured at 2.04 Torr with 80% uncertainty.

It is difficult to ascertain the quality of both the *ab initio* dipole moment surface and the experimental relative intensities, given the observations described above. Also, there are differences observed in identical bands across separate pressure measure-

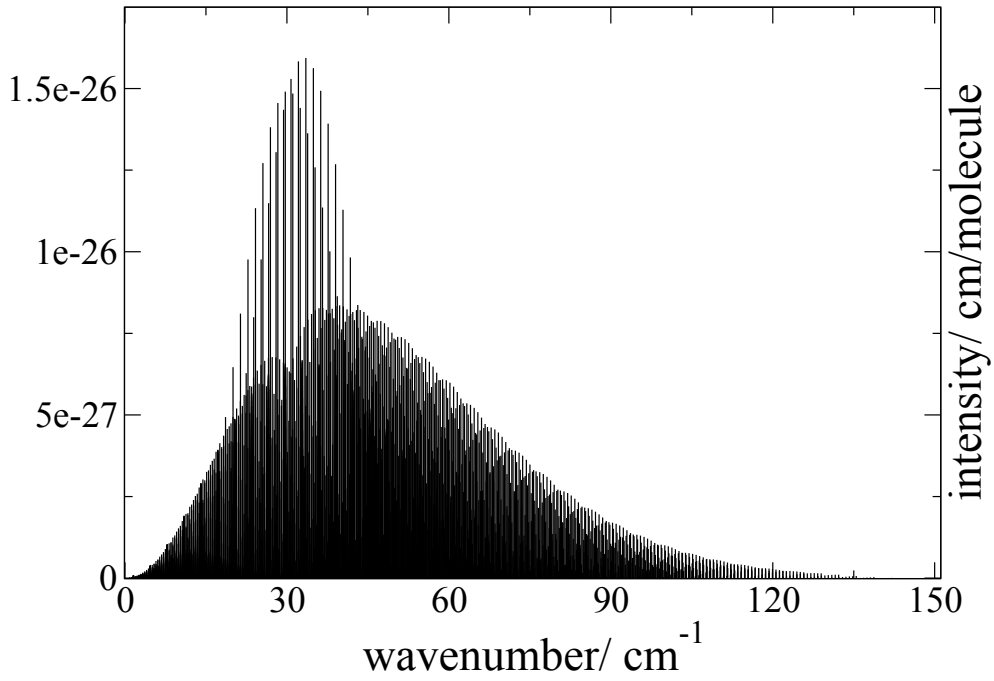

Figure 4.3: Rotational absorption band computed for T = 298.15K, complete up to J=85.

ments. In the case of the $2\nu_4^{(l_4=0)}$ - ν_4 band the comparison with the 2.04 Torr experiment yields a better value for $S_{\text{red}}^{\text{calc}}/S^{\text{exp}}$ than for the 0.409 Torr measurement. This could probably be due to the number of lines available at 0.409 Torr being too low (see also Figure 4.4).

Nevertheless, based on these numbers it is possible to place a lower estimate on the quality uncertainty for the *ab initio* intensities for each band; for example the experiment and theory for the ν_2 and ν_4 bands agree at least to within 3% for both pressure measurements, while it is only 13% for $2\nu_2$ - ν_2, and between 17% -18% for the remaining bands.

In conclusion, the complete room-temperature *ab initio* line list for SO_3 contains 174 674 257 lines with transition energies, Einstein coefficients $A(f \leftarrow i)$, and absorption intensities estimated for T = 298.15 K. Additionally, a list of 10 878 experimental transitions have been associated with computed absolute intensities obtained for T = 296 K and have been included in the latest edition of HITRAN [75]. With the computation of the more accurate hot line list to follow, both this lists will be updated.

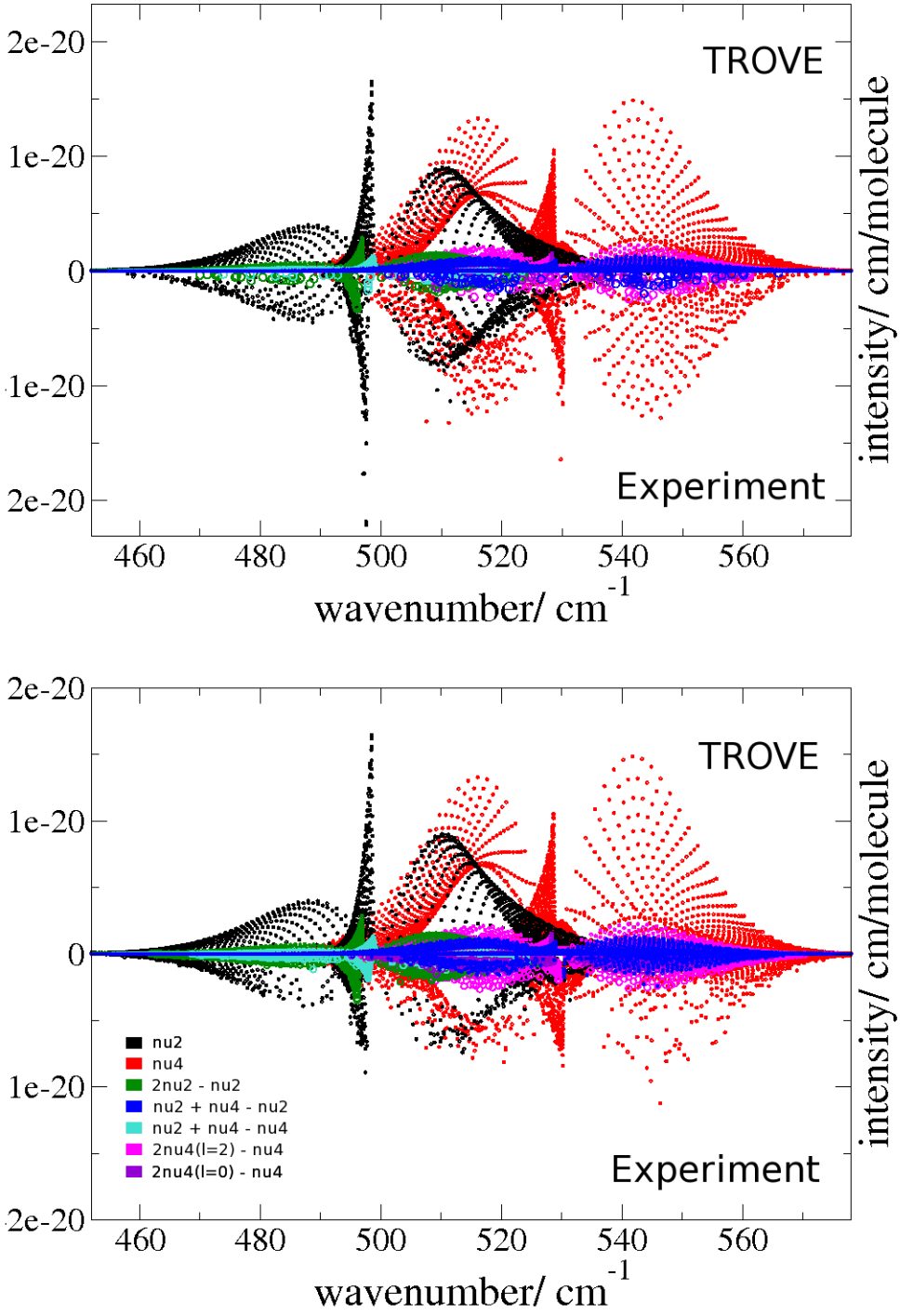

Figure 4.4: Comparison plot of TROVE results and the bands of interest measured between 405 - 708 cm^{-1} by Maki *et al.*, at 0.409 Torr (above) and 2.04 Torr (below). Points are enlarged in some cases for clarity.

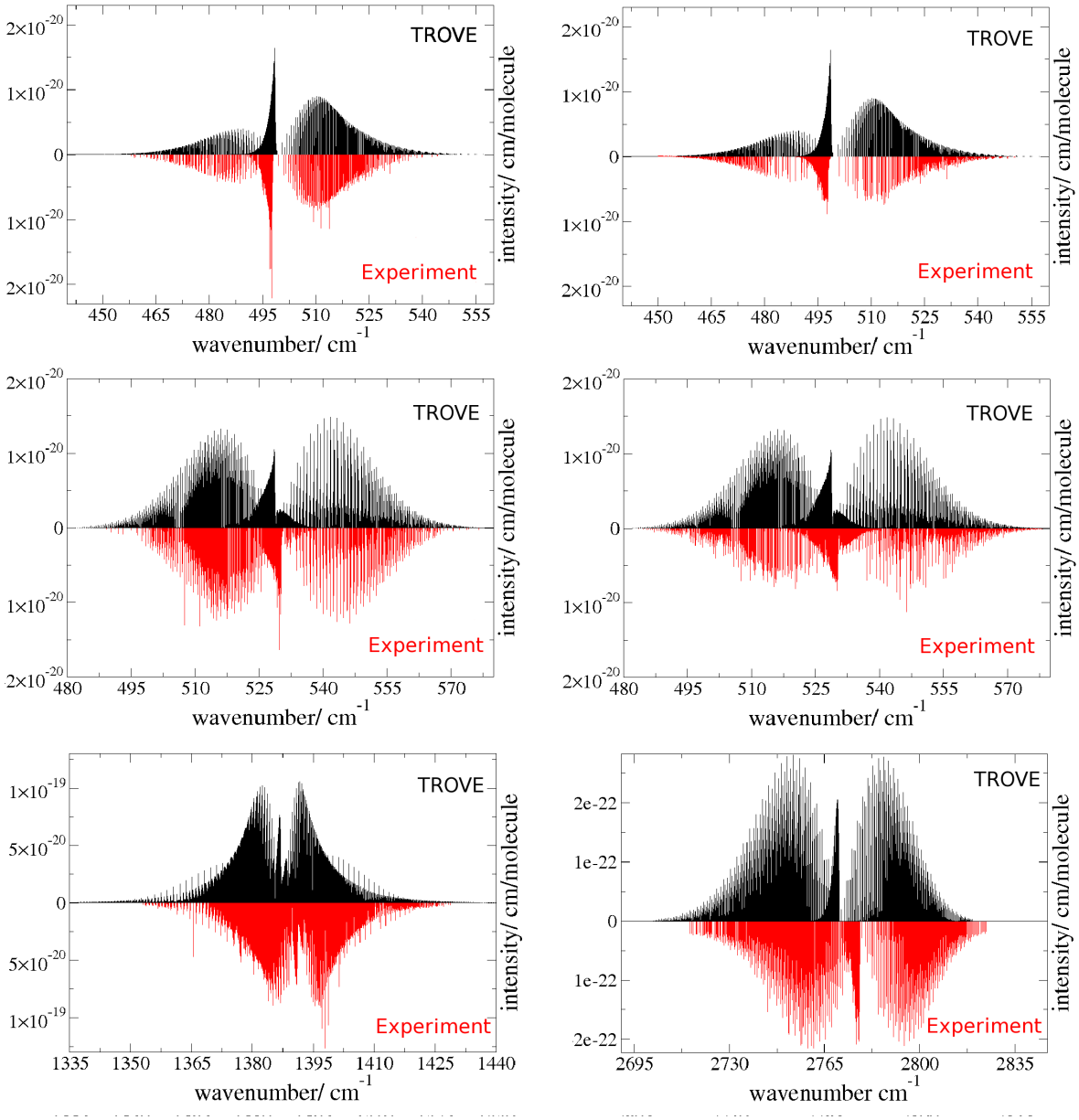

Figure 4.5: Fundamental band comparisons between this work (T = 298.15K, complete up to J=85) and Maki *et al.* [104]. The top and middle panels show comparisons for the ν_2 and ν_4 bands, respectively, at 0.409 Torr (left) and 2.04 Torr (right). The bottom panel shows a comparison for the ν_3 band at 0.16 Torr (left) and the $2\nu_3$ band at 4.99 Torr (right).

Chapter 5

The 'Forbidden' Rotational Spectrum of $^{32}S^{16}O_3$ and Formation of Sixfold Rotational Energy-Level Clusters in its Vibrational Ground State

The calculated transitions described in Chapter 4 include the 'forbidden' rotational spectrum for $^{32}S^{16}O_3$. Induced by centrifugal distortions of bond lengths and angles, these transition lines exhibit an interesting structure. Experimentally, only a handful of microwave lines have been measured [106]. As well as these lines being few in number, no intensity values are available for them. In this chapter an investigation of the behaviour of the rotating SO_3 molecule is made from a theoretical perspective.

As well as interpreting the rotational spectrum, the phenomenon of near-degenerate energy cluster formation in the highly excited rotational states of SO_3 is discussed. The analysis of the rotational motion of SO_3 provides an opportunity to characterize, for the first time, such energy cluster formation for a planar tetratomic molecule containing no hydrogen.

Dorney and Watson [203] were the first to explain the formation of quasi-degenerate energy levels for spherical top molecules using classical arguments followed by a series of works based on the classical description of the clustering phenomena [204–219].

Most of these descriptions involve the motion of the molecule about a 'localiza-

tion' axis at high rotational excitation [203,207,211,214]. The clustering phenomenon was further explored in great detail using a formalism based on the rotational energy surface (RES) [220]. In this approach, a RES is introduced to describe the classical rotational energy of the molecule as a function of the classical angular momentum vector, and equivalent stationary points on this surface, so-called relative equilibria, in particular equivalent maxima directly associated with the patterns of rotational energy clusters. For a molecule with structural symmetry (resulting from identical nuclei arranged symmetrically at equilibrium), the rotational energy clusters are associated with the emergence of symmetrically equivalent localization axes recognizable on the RES.

Typical molecules exhibiting rotational energy clusters are of the type XY_N with a heavy central atom X, Y being H or D, and the X-Y_i bonds being nearly orthogonal [167, 221], which are often referred to as local mode molecules [167, 221]. The formation of rotational or rovibrational energy-level clusters of this type of molecules was studied for triatomic XH_2 molecules (see, for example, Refs. [167,214,221–223]), pyramidal XH_3 molecules, including PH_3 [167, 218, 224, 225], BiH_3, SbH_3 [194] and their deuterated isotopologues [226]. Molecules with heavier Y-atoms can also form rotational energy clusters. The existence of the clustering was predicted for the molecules CF_4 [227–229], SF_6 [204–206], P_4 [216] and experimentally confirmed for SF_6 [230].

For PH_3, BiH_3 and SbH_3, the energy cluster formation was investigated theoretically for $J \leq 70$. SO_3 is not a local mode molecule [167, 221], but, as will be shown, for SO_3 a pattern of energy clusters is predicted in the present work. However, the clustering behaviour does not manifest itself until the rotational excitation reaches the J range between 100 and 250. SO_3 is also much heavier than the XH_3 molecules, it has much smaller rotational energy spacings due to the small rotational constant and, in consequence, states with much higher J values will be populated at a given temperature. This makes the clustering states of SO_3 potentially accessible experimentally at temperatures of a few hundred °C. Another potential opportunity to reach the required rotational excitations is the rotational centrifuge technique, which allows one to climb the rotational ladder up to the dissociation limit [231].

The present chapter deals with two aspects of the SO_3 rotational motion. First, rovibrational energies, wavefunctions and transition intensities produced in the preliminary room-temperature line list study for SO_3 (Chapter 4) are used to simulate

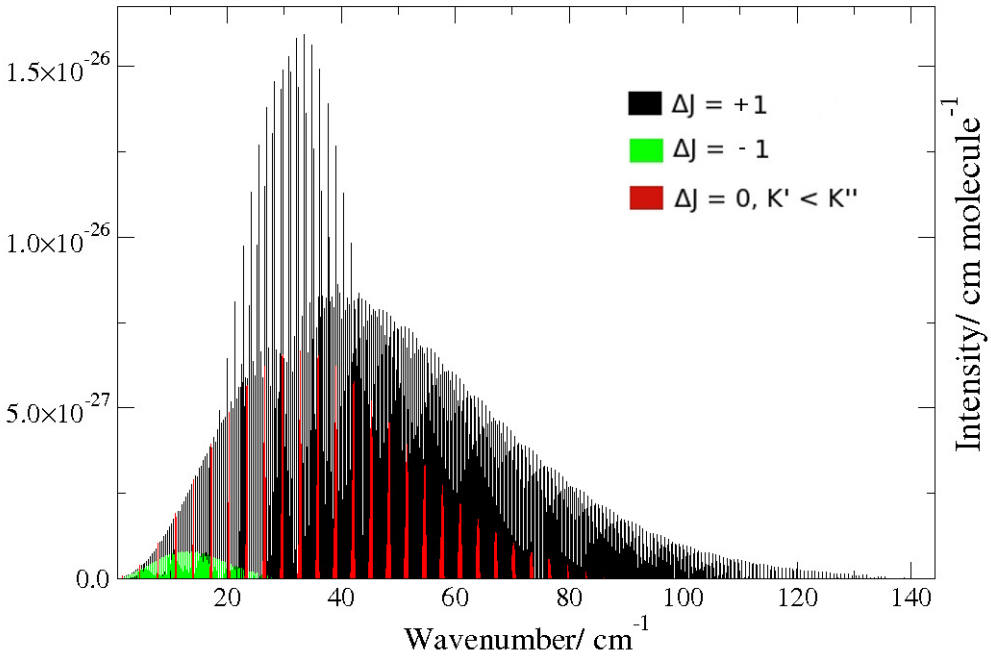

Figure 5.1: The microwave spectrum of $^{32}S^{16}O_3$, simulated theoretically at $T = 298.15$K. Transitions between states with $J \leq 85$ are included; they are color-coded in the diagram to indicate the various transition types.

the room-temperature microwave (rotational) spectrum, and to analyse its structure. The *ab initio* intensities have been combined with line positions obtained from an effective, Watsonian-type rotational Hamiltonian, whose parameter values have been determined in least-squares fittings to experimentally observed transition frequencies [99–105]. This is expected to produce an accurate rotational line list for SO_3, which is recommended for inclusion in the HITRAN [75] and other spectroscopic databases [76, 124]. Secondly, for the first time, the formation of rotational energy level clusters for the planar $\mathscr{D}_{3h}$(M)-type molecule SO_3 is described both classically and quantum-mechanically.

5.1 The Theoretical Microwave Spectrum

Figure 5.1 shows the simulated microwave spectrum for $^{32}S^{16}O_3$, computed with TROVE at $T = 298.15$ K. The spectrum consists of 3439 transitions between all rotational states with $J \leq 85$. The transitions are subject to the selection rules $\Delta J = J' - J'' = 0, \pm 1$ $(J'' + J' \geq 1)$, and symmetry selection rules $A_1' \leftrightarrow A_1''$. The intensity of a transition between two states is defined as in Equations 4.11 and 4.12 of

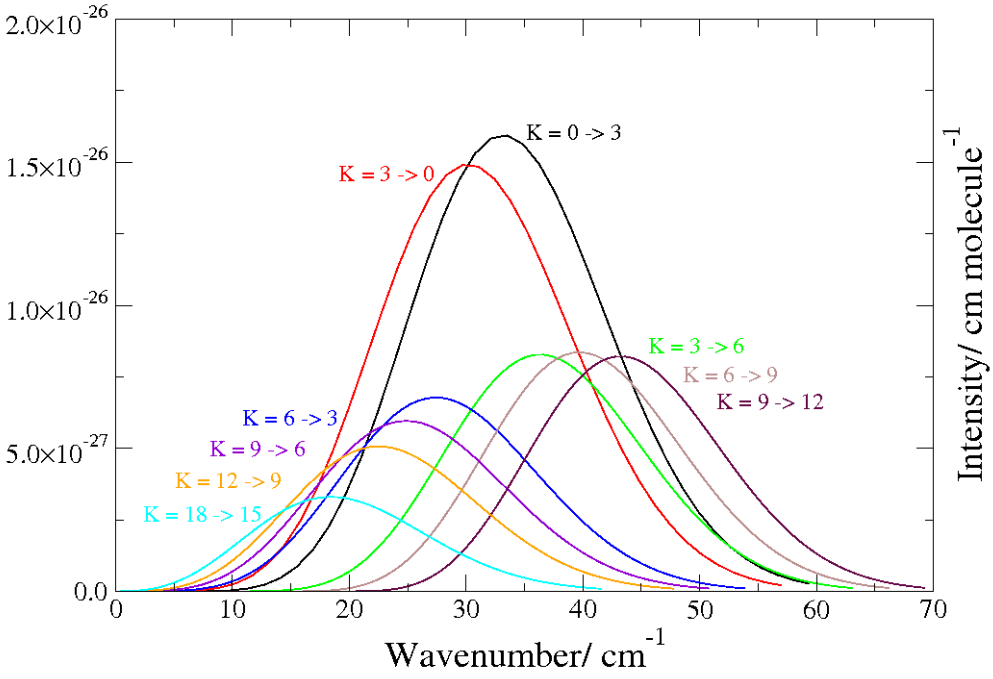

Figure 5.2: The strongest individual K sub-bands of the R-branch in the purely rotational spectrum of $^{32}S^{16}O_3$.

Chapter 4.

The colour coding in Fig. 5.1 distinguishes the P-, Q-, and R-branches in the spectrum. It can be seen that in each branch the intensities depend on J in a characteristic manner. Figure 5.2 expands the wavenumber scale for the R-branch and the structures of the individual K sub-bands are indicated. The intensities of the K sub-bands decrease gradually as the K-values involved increase. SO_3 has $\mathscr{D}_{3h}$(M) symmetry at equilibrium and, in consequence, no 'permanent' dipole moment. Therefore, the transitions in Figure 5.1 are 'forbidden', that is, they are induced by the rotation-vibration interaction as explained, for example, in Section 14.1.14 of Ref. [232]. The transitions satisfy the selection rules $A_1' \leftrightarrow A_1''$ mentioned above and states of rovibrational symmetry A_1' or A_1'' occur when the value of the quantum number K (which is not a 'good', rigorously defined quantum number, but which can nevertheless be conveniently used for labelling the SO_3 rotational energies) is a multiple of 3: $K = 0, 3, 6, \ldots$. Figure 5.2 shows that the strongest sub-bands have $K = 0 \leftarrow 3$ and $K = 3 \leftarrow 0$. In general, strong sub-bands satisfy $\Delta K = \pm 3$ and the band intensity decreases as $K \rightarrow J$.

In the following, an analysis of the wavefunctions for the states involved in the

'forbidden' rotational transitions of SO_3 is presented, and their non-vanishing line strengths are discussed as a consequence of intensity stealing [232].

The wavefunctions are computed as linear combinations of the vibrational and rotational basis functions $\phi_{\lambda}^{(J=0)}|J,K,\tau_{\rm rot}\rangle$ discussed previously in Chapter 2. Thus the wavefunctions are given as

$$\Psi_n^{J,\Gamma} = \sum_{K,\tau_{\rm rot}}^{J} \sum_{\lambda} C_{\lambda,K,\tau_{\rm rot}}^{(n,J,\Gamma)} \phi_{\lambda}^{(J=0)} |J,K,\tau_{\rm rot}\rangle. \tag{5.1}$$

Here, the $C_{\lambda,K}^{(n,J,\Gamma)}$ are expansion coefficients, Γ is the irreducible representation of the $\mathscr{D}_{3\rm h}$(M) molecular symmetry group generated by $\Psi_n^{J,\Gamma}$, n is a running index that distinguishes eigenfunctions that have the same J and Γ. J, K, and $\tau_{\rm rot}$ are rotational quantum numbers, and λ specifies a vibrational state.

The vibrational basis functions $\phi_{\lambda}^{(J=0)}$ that enter into Eq. (5.1) are expressed as linear combinations of products of 1D primitive basis functions ϕ_{n_i}. Each contracted basis function $\phi_{\lambda}^{(J=0)}|J,K,\tau_{\rm rot}\rangle$ is labelled by the quantum number values for the primitive basis functions with the largest contribution (in terms of the square of the expansion coefficient) to this contracted basis function. The quantum numbers labelling the primitive vibrational functions are the principal quantum numbers of the individual vibrational modes which can be conveniently mapped onto the conventional harmonic-oscillator quantum number v_i traditionally used to label vibrational states. In practice, the vibrational labels $\lambda = (v_1, v_2, v_3, v_4)$ are used for the $\phi_{\lambda}^{(J=0)}$ basis functions where, in the standard manner, v_i is associated with the normal mode ν_i, $i = 1, 2, 3, 4$. That is, the vibrational ground state has $\lambda = (0,0,0,0)$, the ν_1 state has $\lambda = (1,0,0,0)$, and so on.

Since purely rotational states are of interest here, it is expected that in Eq. (5.1), the largest contribution to the ground vibrational state $\Psi_n^{J,\Gamma}$ originates in a $\lambda = (0,0,0,0)$ basis function. That is, the largest value of $|C_{\lambda,K,\tau_{\rm rot}}^{(n,J,\Gamma)}|^2$ is obtained for $\lambda = (0,0,0,0)$. Indeed, this is what is observed in the calculation. However, the expansion in Eq. (5.1) is not generally dominated by this one term only, there will normally be a number of other terms that also have significant contributions.

To understand better the significance of the contributions from the contracted basis functions, which in themselves are truncated linear combinations of primitive basis

functions, the vibrational reduced density is defined as

$$A_{\lambda}^{n,J,\Gamma} = \sum_{K,\tau_{\rm rot}^{J}} |C_{\lambda,K,\tau_{\rm rot}}^{(n,J,\Gamma)}|^2. \tag{5.2}$$

The summation is over K and $\tau_{\rm rot}$ for all expansion-coefficient squares associated with a particular vibrational basis function $\phi_{\lambda}^{(J=0)}$; the individual terms in the sum originate in the basis functions resulting from the combination of the particular function $\phi_{\lambda}^{(J=0)}$ with different rotational basis functions $|J,K,\tau_{\rm rot}\rangle$. The resulting quantities $A_{\lambda}^{n,J,\Gamma}$ indicate how the eigenfunction $\Psi_n^{J,\Gamma}$ is partitioned between the vibrational basis functions $\phi_{\lambda}^{(J=0)}$.

Figures 5.3(a) and 5.3(b) show how various $A_{\lambda}^{n,J,\Gamma}$ vary with K_{lc}, which specifies the K corresponding to the largest contribution to $A_{\lambda}^{n,J,\Gamma}$. Each wavefunction $\Psi_n^{(J,\Gamma)}$ is found to have has dominant contributions from the $\phi_{\lambda}^{(J=0,A_1')}$ basis functions with $\lambda = (0,0,0,0)$ and thus associated with the vibrational ground state. However, there are other significant contributions originating in $\phi_{\lambda}^{(J=0)}$ basis functions with $\lambda = (0,1,0,0)$ and $(0,0,0,1)$ associated with the ν_2 and ν_4 bending vibrational modes, respectively. Thus, even in the vibrational state conventionally labelled as the ground state, for purely rotational excitation there is a non-negligible probability of finding the molecule with the ν_2 or ν_4 vibrational mode excited by one quantum.

The contributions to the eigenfunctions from the ν_2 and ν_4 basis states are highest at $K = 0$ and drop as $K \to J$ (see Figs. 5.3(a) and 5.3(b)). This can be straightforwardly explained classically. For a given J-value, the rotational eigenstates of SO_3 that are conventionally label by $K = J$ and $K = 0$ correspond to the two limiting cases of the molecule rotating about the z-axis (which, in configurations of $\mathscr{C}_{3v}$ geometrical symmetry, coincides with the C_3 symmetry axis) and about axes perpendicular to the z axis, respectively. A rotation about the z axis preserves the high $\mathscr{D}_{3h}$(M) geometrical symmetry that the molecule has at equilibrium and the dipole moment vanishes at $\mathscr{D}_{3h}$(M)-symmetry configurations, also for $K \to J$. For $K = 0$, the situation is quite different. The angular momentum vector of length $\sqrt{J(J+1)}\hbar$ now results from rotation about axes perpendicular to the z axis and this, in turn, gives rise to centrifugal forces that deform the molecule away from $\mathscr{D}_{3h}$(M) geometrical symmetry. These deformations can be described as excitations of the deformation bends ν_2 and ν_4 and, in the resulting

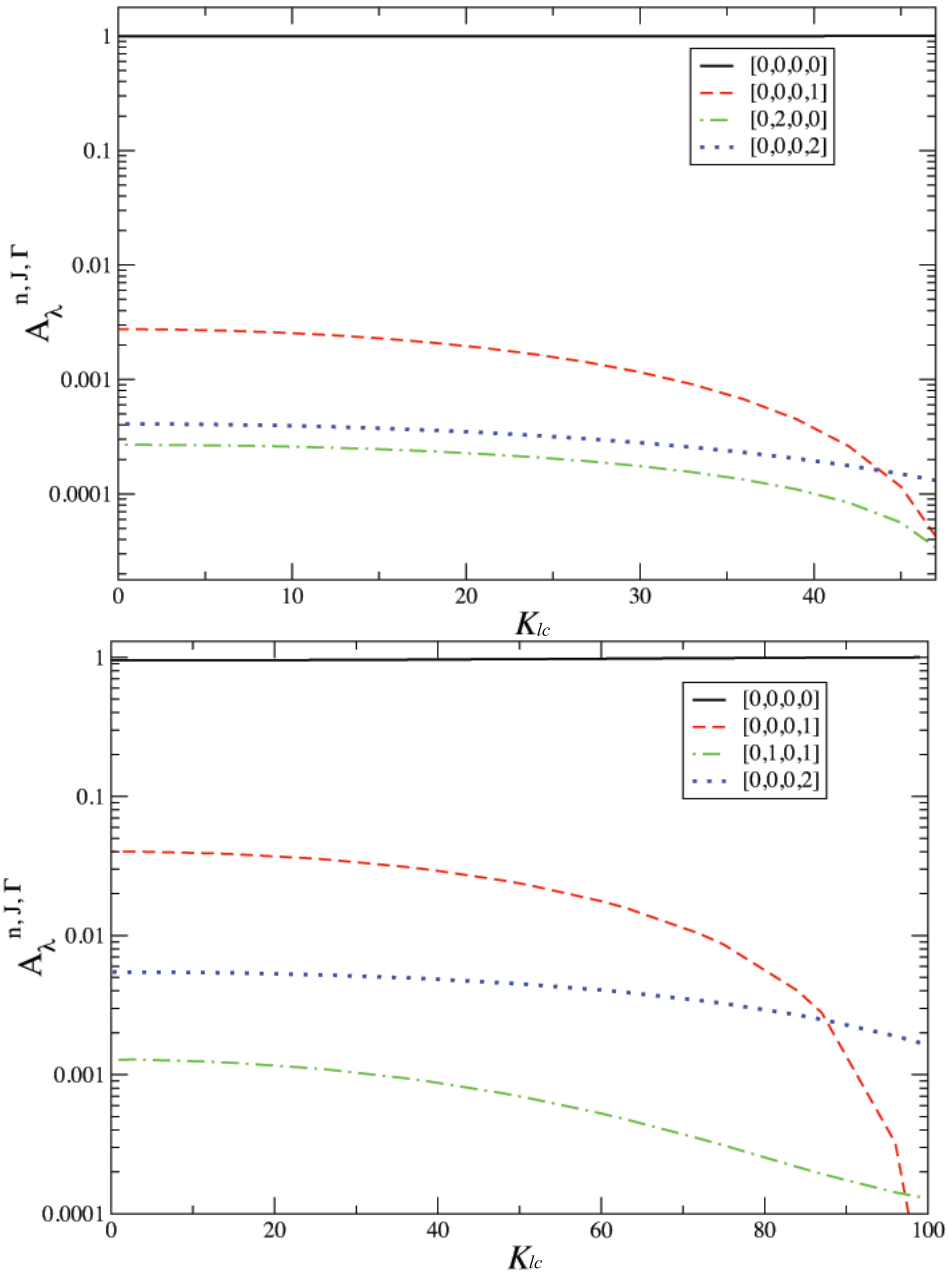

Figure 5.3: Fig. (a) and (b), $J = 50$ and $J = 100$, respectively: The vibrational reduced density values $A_\lambda^{n,J,\Gamma}$ (Eq. (5.2)) for specific pure rotational states $\Psi^{J,\Gamma}_{[0,0,0,0],K,\tau_{\text{rot}}}$ with (a) $J = 50$ and $0 \leq K \leq 50$ and (b) $J = 100$ and $0 \leq K \leq 100$, respectively, of $^{32}S^{16}O_3$, plotted against the (approximate) quantum number K for four largest contracted basis set contributions $\lambda = [v_1, v_2, v_3, v_4]$.

molecular configurations, the dipole moment has non-vanishing instantaneous values.

This explains the form of the rotational band in Fig. 5.1: the strongest transitions involve the states with the smallest K-values. The transitions in the two bands with $K = 0 \leftarrow 3$ and $K = 3 \leftarrow 0$, respectively, have intensities (and Einstein coefficients) about twice as strong as the transitions in the bands involving states with $K \neq 0$ only. Quantum mechanically, this is explained by the special-case nature of the $K = 0$ wavefunctions giving rise to an 'extraneous' factor of 2 in the expression for the line strength (see, for example, Eq. (12.33) of Ref. [154]). A more detailed analysis of the rotational eigenfunctions obtained at very high J-values is presented below in connection with the discussion of the rotational-cluster-formation phenomenon.

The nonvanishing intensity of the SO_3 rotational spectrum is explained here in terms of intensity stealing. Alternatively but equivalently, it can be viewed as originating in asymmetric centrifugal distortions of the molecule leading to the formation of a small, non-vanishing dipole moment in the vibrational ground state, see, for example the prediction of a pure rotational spectrum of H_3^+ [233, 234].

5.2 A Hybrid, Empirical Ab Initio Rotational Line List for $^{32}S^{16}O_3$

In the previously generated list of transitions for $^{32}S^{16}O_3$, which were included in the 2012 edition of the HITRAN database [75], available experimental transitions [99–105] were taken and compared to entries in the computed room-temperature line list, and their relative intensities were adapted to the absolute scale [170] computed *ab initio* using TROVE (see Chapter 4). The line positions of 25 pure rotational transitions were measured by Meyer *et al.* [106] and the synthesised HITRAN line list for $^{32}S^{16}O_3$ included these 25 experimental microwave lines augmented with the theoretical values computed for the transition intensities.

Here the room-temperature line list is extended to include additional rotational transitions. For this, the ground state rotational constants and centrifugal distortion

parameters for $^{32}S^{16}O_3$ of [104] are used in conjunction with the expression

$$\begin{aligned} F(J,K) &= B_0 J(J+1) + (C_0 - B_0)K^2 - D_0^J[J(J+1)]^2 - \\ &- D_0^{JK} J(J+1)K^2 - D_0^K K^4 + H_0^J[J(J+1)]^3 + \\ &+ H_0^{JK}[J(J+1)]^2 K^2 + H_0^{KJ} J(J+1)K^4 + \\ &+ H_0^K K^6 \pm \delta_{3K}\Delta_0[J(J+1)][J(J+1)-2][J(J+1)-6] \end{aligned} \quad (5.3)$$

to generate rotational energies for J up to 85, using a simple script. These constants have already been used in the analysis of the spectra presented in Refs. [99–105].

To make sure that the extrapolation to high-J states is reasonable, a comparison between the energy values determined from Eq. (5.3) and the corresponding transition wavenumbers generated by TROVE has been made. A very good agreement is observed, with a root-mean-square difference of only 0.0167 cm^{-1} for all rotational transitions involving states with $J \leq 85$.

Table 5.1 shows examples of comparisons between rotational term values in the vibrational ground state of $^{32}S^{16}O_3$ calculated from Eq. (5.3) and values obtained variationally with TROVE. The 25 lines from the previous HITRAN-suitable list were taken from Meyer *et al.* [106]. These transitions are replicated by Eq. (5.3) to within the precision quoted in the updated list. It is expected that the values from Eq. (5.3) are more accurate than the TROVE values, so the line positions generated with this equation are used in combination with their attributed absorption intensities (Einstein coefficients) determined from the *ab initio* dipole moment surface. These new 3414 transitions for $^{32}S^{16}O_3$ should make a good addition to HITRAN and other spectroscopic databases; they constitute a significant extension on the 25 microwave transitions presently included in HITRAN.

5.3 Rotational Energy Level Cluster Formation in High-J States

To investigate the possible rotational-energy cluster formation in SO_3, rovibrational term values of $^{32}S^{16}O_3$ for $J \leq 250$ have been computed. The present work is focused on rovibrational states in the vibrational ground state, and so the calculation aims at

Table 5.1: Comparison of rotational term values for $^{32}S^{16}O_3$ obtained by inserting the experimentally derived, effective rotational constant values from [104]) in Eq. (5.3) ('Obs.') with theoretical TROVE values from the present work ('Calc.').

J	K	Obs./ cm^{-1}	Calc./ cm^{-1}
5	3	8.885	8.886
10	9	24.200	24.203
10	0	38.336	38.342
20	18	89.825	89.837
20	0	146.333	146.355
50	48	486.469	486.514
50	0	886.775	886.897
85	84	1315.569	1315.283
85	3	2529.972	2530.258

obtaining accurate results for these states only. Because of this, it is possible to reduce the size of the vibrational basis with minimal loss in accuracy: only basis functions satisfying $P \leq P_{\max} = 8$ are employed in the polyad truncation scheme (see Chapters 2 and 4).

Rotational term values for SO_3 are denoted as $E_{J,K}$ and the reduced energy is defined as $\Delta E_{J,K} = E_{J,K} - E_{J,K}^{\max}$ where $E_{J,K}^{\max}$ is the maximum energy found in a given J multiplet of the vibrational ground state (see, for example, Refs. [194, 224, 225, 235]). The rotational-energy-cluster formation can be visualised by plotting the reduced energies $\Delta E_{J,K}$ against J in Fig. 5.4(a) for all states including those forbidden by the nuclear spin statistics. The figure shows that for $J > 150$, energy-level clusters form for $^{32}S^{16}O_3$ in a manner similar to that previously described for the pyramidal molecules PH_3, BiH_3 and SbH_3 [194, 224, 225, 235] at significantly lower J values.

The four highest energies in each J manifold cluster together. In conventional symmetric-top notation, these highest four energy levels, taken in descending order, have $K = |k| = 0$, 1, 2, and 3, respectively, and the $\mathscr{D}_{3h}$(M) symmetries A_1', E'', E', and A_2'' (A_2', E'', E', and A_1'') for even (odd) J. Since the states of E symmetry are doubly degenerate, the four energies define a near-degenerate six-fold cluster. Only one of these fours levels for each J has symmetry allowed by the nuclear statistics (A_1' or A_1''), see Fig. 5.4(a).

Figure 5.4(b) compares the reduced energy values obtained with TROVE with values determined from Eq. (5.3), using the experimentally derived values of the required rotation-vibration parameters from [104]. Although the reduced energies obtained from

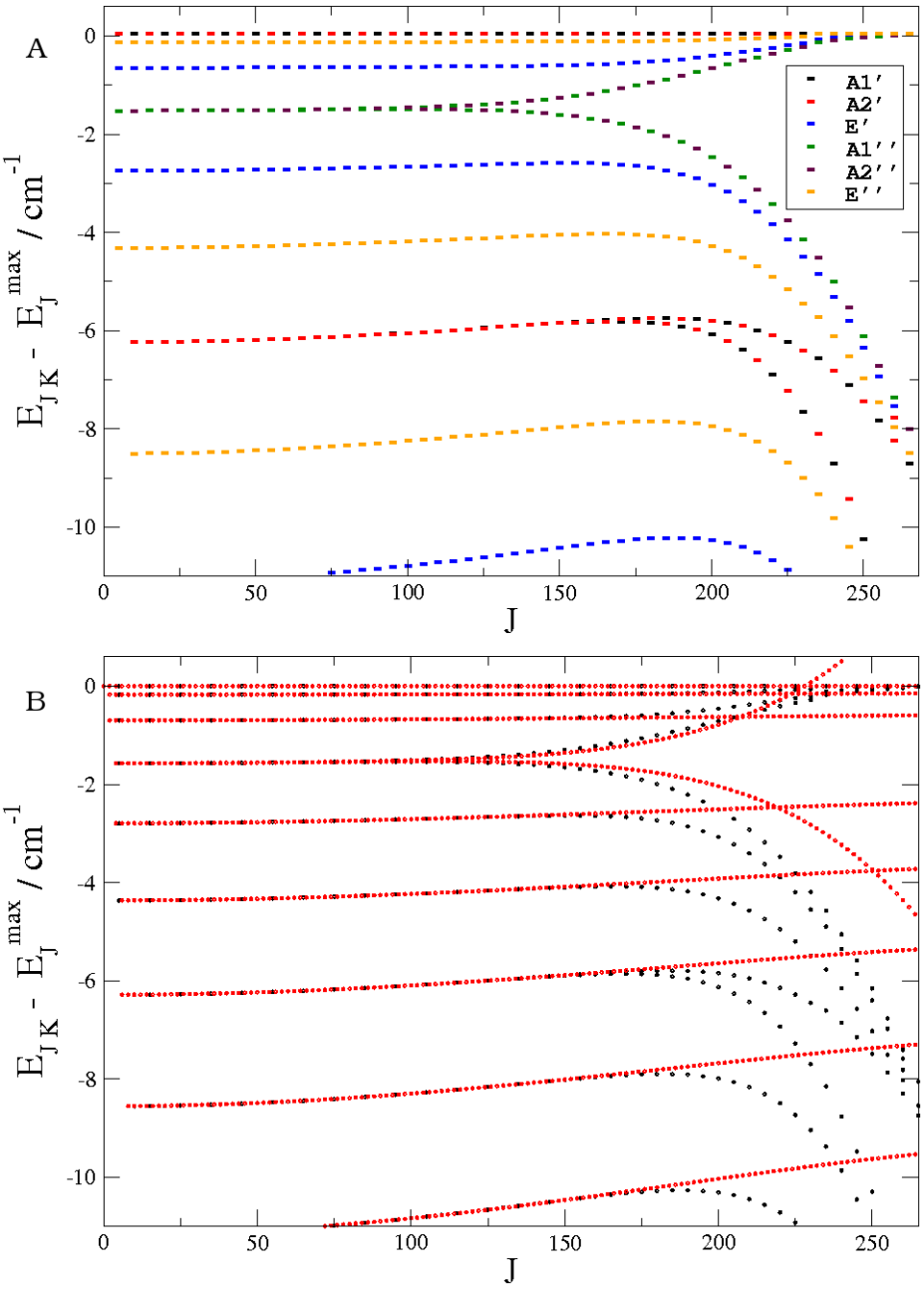

Figure 5.4: Fig. (a), Reduced energies $\Delta E_{J,K}$ plotted against J for $^{32}S^{16}O_3$. The $\Delta E_{J,K}$-values are calculated with TROVE. The symmetry of the $E_{J,K}$ levels is indicated by color. We note that only A_1' (black) and A_1'' (green) are allowed by Bose-Einstein statistics. Fig (b). Reduced energies $\Delta E_{J,K}$ plotted against J for $^{32}S^{16}O_3$. Reduced energy values obtained with TROVE (black) are compared to experimentally derived values (red) from Ref. [104].

Maki *et al.* Eq. (5.3) do not exhibit a 6-fold cluster formation at the high J limit, there is a clear tendency in the shape of the $K = 3$ curve to approach the topmost, $K = 0$ energy levels. However all other rotational states are completely misplaced indicating the deficiency of the simplified model of Eq. (5.3) for describing the energy levels with very high J.

The expansion (5.3) is truncated after the 6th order terms and it is expected that higher-order terms become important at $J > 150$. In consequence, Eq. (5.3) is not suitable for extrapolation to states with so high J-values, and the TROVE results are taken to be more reliable.

5.3.1 Origin of the Cluster States

It is accepted and has been described in detail that the rotational clustering is associated with the symmetry breaking of rotational energy surfaces and corresponding stationary points [217]. For the XY_3 type pyramidal molecules the formation of six-fold rotational-energy clusters can be explained in terms of classical arguments that are outlined briefly here. In the cluster states, the rotation of XY_3 can be thought of as taking place around one of three possible, so-called localisation axes, which represent the stationary points, or relative equilibria of the corresponding RES [194, 224, 235]. For pyramidal XY_3 molecules these three symmetrically equivalent axes are of the C_2 type and nearly coincide with the three X-H_i bonds of the XH_3 molecule, where H_i is the proton labelled $i = 1$, 2, or 3 [194,224]. For each of the three localisation axes there is the possibility of clockwise or anticlockwise rotation and so six equivalent physical situations emerge, all six having approximately the same energy. The stationary points are the six maxima on the RES [235], which constitute a six-fold near-degenerate energy cluster. The cluster states are stabilised by large energy barriers separating each of the six equivalent situations just described.

Quantum mechanically to describe one of these situations it is customary to introduce the concept of a *primitive cluster state* (PCS) function (see, for example, [221, 224, 236]). When the PCSs are separated by energy barriers so high as to effectively prevent tunneling between the states, the six equivalent PCS situations become degenerate and the manifold of wavefunctions associated with their common energy can be viewed as a six-dimensional space where the PCS, or coherent wave-

functions serve as a convenient basis set. Jensen and Kozin [236] constructed the PCS wavefunctions of an XH_2 molecule using the symmetry properties of the corresponding eigenstates. Yurchenko, Theil, Patchkovskii and Jensen [224] obtained the PCS wavefunctions for an XH_3 molecule by diagonalising a matrix representing the angular momentum. This latter approach is followed below in constructing and visualising the (quantum-mechanical) PCS wavefunctions for SO_3, a planar XY_3 molecule.

Both the classical and quantum mechanical approaches are outlined in further detail below.

5.3.2 Classical Analysis in Terms of the Rotational Energy Surface

Classically, the rotational cluster formation can be understood in terms of the topology of the rotational energy surface [208, 210, 215].

Qualitatively speaking, the concept of the rotational energy surface (RES) is analogous to that of a potential energy surface (PES); the adiabatic rotational energy within an isolated electronic and vibrational state is represented by a continuous 2-dimensional surface embedded in 3-dimensional space. For a given value of J, the RES is obtained as a radial plot of the function $E_J(\theta,\phi)$. This function yields the classical rotational energy in terms of the two angles (θ,ϕ) defining the orientation of the classical angular momentum vector in the molecule-fixed axis system *xyz*. The RES serves to track the motion of the laboratory fixed angular momentum vector $\hat{J}$ in the molecule-fixed frame, i.e. each point on the RES represents a direction of $\hat{J}$ in the molecule-fixed frame, and the magnitude at each point is taken to be the adiabatic rotational energy (in arbitrary units) [208].

The trajectory of $\hat{J}$ on the RES is perhaps best understood also as an analogy, to the classical trajectory along a PES. The trajectory along the potential is confined to a path bound by classical turning points, which are defined as the intersection of the PES with a 'constant energy surface' (CES). In the case of a RES, the intersecting CES is spherical, since the energy plot is radial. Therefore, for the purely kinetic rotational energy, the classical trajectory of $\hat{J}$ lies upon the contour line defined by the intersection of the RES with a particular CES. Since a particular RES corresponds to a certain J value, the different contour lines on the RES defined by intersections of multiple CESs serve to visualise the fine structure of the rotational dynamics. Figure 5.5 [208] shows

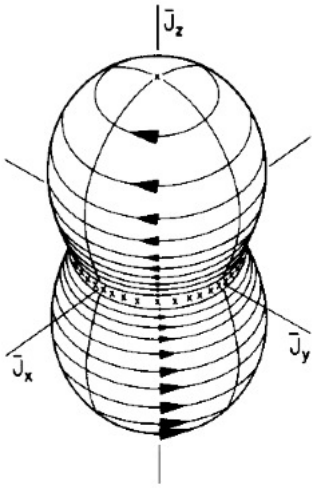

Figure 5.5: Rotational energy surface for a rigid, prolate symmetric top [208]. The contour lines of RES and CES intersection define the classically allowed trajectories of the $\hat{J}$ vector.

an example of a RES for an arbitrary rigid, prolate symmetric top, where such contour lines can be seen.

Thus, the *xyz* components of the classical angular momentum vector, (J_x, J_y, J_z), are given by

$$J_x = \sqrt{J(J+1)}\hbar \sin\theta \cos\phi \tag{5.4}$$

$$J_y = \sqrt{J(J+1)}\hbar \sin\theta \sin\phi \tag{5.5}$$

$$J_z = \sqrt{J(J+1)}\hbar \cos\theta\,, \tag{5.6}$$

where $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$. In this semi-classical approach the angular momentum vector is assumed to have its 'quantum-mechanical' length of $\sqrt{J(J+1)}\hbar$.

In order to construct the RES a perturbation theory-free approach [224,235,236] is followed to obtain a classical Hamiltonian function H_{rv} by applying the classical limit to the quantum-mechanical Hamiltonian operator $\hat{H}$, i.e. to an "*ab initio* Hamiltonian" as was proposed by [210]. The method is outlined extensively elsewhere [235]; a brief outline is given here.

The classical Hamiltonian is introduced as a function of J_α, as well as the generalised internal (vibrational) coordinates ξ_n ($n = 1, 2, \ldots, 6$) and their conjugate momenta p_n:

$$H_r = H_{r-v}(\hat{\xi}_e, \hat{p}_e, \hat{J}) \tag{5.7}$$

Here, $\hat{\xi}_e$ and $\hat{p}_e$ are obtained as stationary points, by equating the partial derivative of the Hamiltonian function with respect to each of the parameters to zero, respectively. These values essentially represent the instantaneous configuration of the internal molecular coordinates as a function of the J_α. That is, for each orientation of the angular momentum vector $\hat{J}$, there exists an 'optimal' geometry of the molecule, i.e. a distortion introduced by rotation-vibration interactions. This is the essence of the PES Born-Oppenheimer analogy; the "vibrational structure" is determined by the "clamped" rotational configuration.

The generalised vibrational momenta p_n are actually set to zero [214], which is shown to be a good approximation [224, 225]. The rotational energy surface $E_J(\theta, \phi)$ is then generated on a regular grid of angle points θ_m, ϕ_m by setting the vibrational coordinates in $\hat{H}$ to the optimised geometries $\hat{\xi}_e$ and $\hat{p}_e$ for each orientation of the angular momentum defined by the polar and azimuthal angles (θ_m, ϕ_m) and Eqs. (5.4)-(5.6):

$$E_J(\theta_m, \phi_m) = H_{r-v}(\hat{\xi}_e, \hat{p}_e, \theta_m, \phi_m; J). \tag{5.8}$$

Finally, the RES is given by a radial plot of the function

$$E_J^{(\mathrm{RES})}(\theta_m, \phi_m) = E_J(\theta_m, \phi_m) - E_J^{(\mathrm{min})}, \tag{5.9}$$

where $E_J^{(\mathrm{min})}$ is the minimum value of $E_J(\theta_m, \phi_m)$.

Plotting the RES along with the classically allowed trajectories of $\hat{J}$ offers insight into the clustering phenomenon. As J increases, the topology of the RES changes such that a number of maxima emerge upon the surface, which become more pronounced with increasing J (see, for example, ref. [224]). The total degeneracy of a cluster corresponds to the number of distinct classical trajectories with the same energy, located on equivalent parts (maxima) of the RES [208]. It has already been mentioned that when energy clusters form for pyramidal XH_3 molecules, the six maxima on the RES corresponding to the top clusters in a J manifold are located in the directions defined by the local C_2 symmetry axes associated with the plane containing a X–H molecular bond.

For SO_3, however, this semi-classical analysis suggests that the six maxima appear for $\theta = 90^\circ$ and $\phi = 30$, 90, 150, 210 and 330°, respectively, all in the molecular plane.

An example of a RES of SO_3, computed for $J = 200$, is shown in Fig. 5.6. The resulting RES has a doughnut-like shape with six symmetrically equivalent maxima in the *xy* plane (i.e., for $\theta = 90^\circ$). These maxima are difficult to observe in Fig. 5.6 but they are better seen in the top display of Fig. 5.7, which shows the slice of the RES along a path where θ_m is kept constant at 90° and ϕ_m is varied from 0 through 360°.

Each of these directions is perpendicular to one of the S–O bonds and also define the orientations of the localisation axes A. The molecular geometries corresponding to the maxima on the RES are illustrated by the two lower panels of Fig. 5.7. For the molecular geometry associated with the first stationary rotational axis and illustrated in Fig. 5.8, the localisation axis $A = C_2'$ is perpendicular to the S-O_1 bond and the molecular structure has a $\mathscr{C}_{2v}$ symmetry with S-O_2 and S-O_3 elongated and the angle α_1 = ∠(O_2-S-O_3) reduced from 120 to 112.4°. These six maxima are relatively small and separated by less pronounced barriers comparing to the clustering of XH_3 pyramidal molecules.

However their rotational energy-clustering structure (see Fig. 5.4) is very similar, despite these differences in the topology of the RESs of a planar SO_3 molecule from that encountered for pyramidal XY_3 molecules. The validity of this analysis can be tested by a further, complementary quantum mechanical discussion, which is outlined in the next subsection.

5.3.3 Quantum Mechanical Analysis

To obtain an understanding of the rotational motion of the cluster states, a quantum mechanical approach can be taken by considering the wavefunctions obtained from the variational calculations. The aim is to visualise the localisation axes associated with the cluster states [221, 224]. For a given J-value, the Schrödinger equation for the Hamiltonian is solved using TROVE; the Hamiltonian matrix blocks are diagonalised for all $\mathscr{D}_{3h}$(M) symmetries, using the basis set defined by P_{max}=8. The six near-degenerate eigenfunctions $\psi_n^{(J)}$ associated with the 'top' cluster in the appropriate J manifold of the vibrational ground state are selected. The aim here is to somehow visualise these wavefunctions.

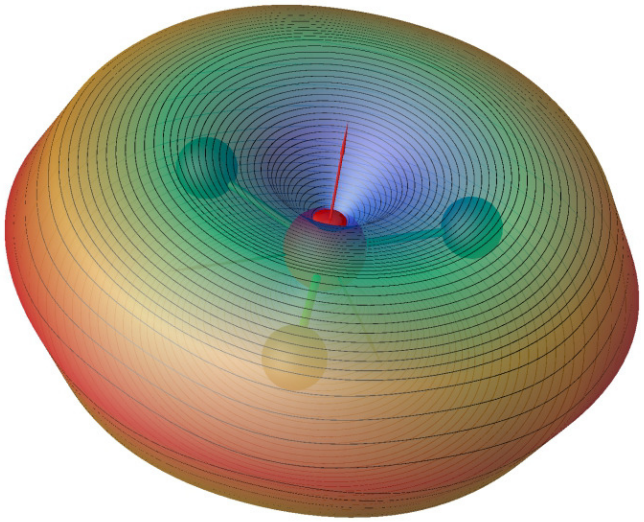

Figure 5.6: The rotational energy surface of SO_3 at $J = 200$.

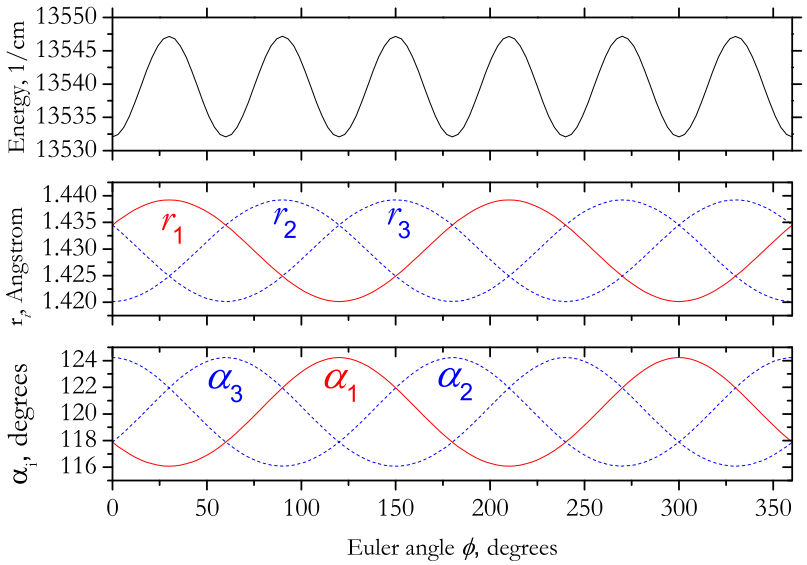

Figure 5.7: Top display: A 1D cut through the $J = 200$ RES of SO_3 for $\theta = 90^\circ$ and $\phi \in [0^\circ, 360^\circ]$. Middle display: The ϕ-dependence (at $\theta = 90^\circ$) of the vibrationally averaged values for the three bond lengths r_1, r_2, r_3. Bottom display: The ϕ-dependence (at $\theta = 90^\circ$) of the vibrationally averaged values for the three inter-bond angles $\alpha_1 = \angle(O_2\text{-}S\text{-}O_3)$, $\alpha_2 = \angle(O_1\text{-}S\text{-}O_3)$, and $\alpha_3 = \angle(O_1\text{-}S\text{-}O_2)$.

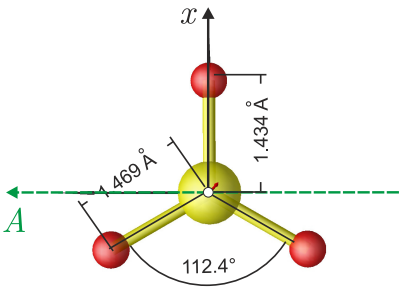

Figure 5.8: The schematic representation of the optimised structure of SO_3 at $J = 200$ corresponding to one of the maxima on the RES.

The states can be visualised by plotting their reduced rotational densities, obtained by integrating the square of corresponding wavefunction $\psi_n^{(J)}$ over all vibrational coordinates and over the Euler angle α:

$$F(\beta,\gamma) = \int_0^{2\pi} \mathrm{d}\alpha \int_{\mathrm{vib}} \mathrm{d}V \left(\psi_n^{(J)}\right)^* \psi_n^{(J)}, \tag{5.10}$$

where $\mathrm{d}V$ is the volume element associated with the vibrational coordinates. This reduction technique reduces the dimensionality to 2; $\left(\psi_n^{(J)}\right)^* \psi_n^{(J)}$ is independent of α, and the integration over the vibrational coordinates serves to isolate the problem to only the rotational motion in the ground state [221, 224].

The explicit form of the rotational functions $D_{J,k,m}(\alpha,\beta,\gamma)$ [232] is used for $m = J$ so that the molecule is aligned with the angular momentum oriented along the space fixed Z axis to the greatest extent allowed by quantum mechanics; this is necessary due to the fact that, even though computed energies are independent of m in the absence of an external magnetic field, the eigenfunctions themselves depend on the value of the quantum number, and therefore require assignment. The consequence of this, classically, is that the Z axis becomes the axis of rotation [224]. The Euler angles β,γ define the orientation of the space-fixed Z axis relative to the molecule-fixed axis system xyz. Therefore, a plot of $F(\beta,\gamma)$ helps one visualise how the molecule orientates itself relative to the Z axis.

In order to explicitly visualise the localisation axes associated with the cluster states, it is necessary to determine the so called "primitive cluster states" $|i\,\mathrm{PCS}\rangle$ associated with each of these axes. Since there is a six-fold near-degeneracy in the cluster

states, six equivalent localisation axes A_i ($i = 1, 2, \ldots, 6$) are assumed to exist, with associated $|i\ \mathrm{PCS}\rangle$, as seen in Ref. [224]. The $\psi_n^{(J)}$ associated with the top six cluster states can be seen to be linear combinations of the $|i\ \mathrm{PCS}\rangle$. To obtain such a PCS a procedure can be followed whereby the six states $\psi_n^{(J)}$ are considered to be fully degenerate, owing to the small differences in their energies; a PCS can therefore be constructed as a linear combination of the $\psi_n^{(J)}$ states, and still be an eigenfunction of the rovibrational Hamiltonian [221, 224].

Once the PCSs are obtained, their corresponding localisation axes A_i can be determined by a procedure outlined in Ref. [224]. In this method the $|i\ \mathrm{PCS}\rangle$ is rotated in space until the largest possible absolute value of the angular momentum projection, $J\hbar$ or $-J\hbar$, is obtained; this orientation is the corresponding localisation axis A_i associated with $|i\ \mathrm{PCS}\rangle$.

The findings of Ref. [224] showed that the $|i\ \mathrm{PCS}\rangle$ occurred in pairs, where two PCSs are associated with a common localisation axis, but each describe 'clockwise' and 'anticlockwise' rotation, respectively. For example, in the two states $|1\ \mathrm{PCS}\rangle$ and $|2\ \mathrm{PCS}\rangle$, the molecule has angular momentum projections of $+m\hbar$ and $-m\hbar$, respectively, with $m \approx J$, onto the localisation axis A_1. Hence the determination of a PCS will automatically obtain its counterpart describing motion in the opposite sense.

In the case of SO_3, the semi-classical argument describes the localisation axes as being in the molecular plane ($\theta_A = 90^\circ$) and therefore only their azimuthal angle ϕ_A must be determined. In agreement with the classical results, the localisation axes are found to be perpendicular to one of the S–O bonds, that is, with $\phi_A = 30$, 90, 150, 210, 270 or 330°. It should be noted that the PCS functions associated with these directions could be also obtained from the six eigenfunctions $\psi_n^{(J)}$ based solely on their symmetries, see, for example, Ref. [236].

Examples of probability distributions $F(\theta, \chi)$ obtained for $\psi = |1\ \mathrm{PCS}\rangle$ at $J =$ 100, 150, 200 and 250 are shown in Fig. 5.9, where the $F(\beta, \gamma)$ functions are imaged onto Bloch spheres of arbitrary radii. The maxima of the $F(\beta, \gamma)$ distributions indicate how the molecule prefers to align itself relative to the space-fixed Z axis and, since $m = J$, to the angular momentum vector which we can think of as being classical in these highly excited rotational states. The classical angular momentum vector is conserved in space and time so that it defines the axis of rotation. The vibrationally averaged

structure, the molecule-fixed axis system *xyz*, and the localisation axis *A* are indicated for the $|1\ \mathrm{PCS}\rangle$ state in Fig. 5.8. The difference between cluster states and non-cluster states is illustrated in Fig. 5.10 which represents the probability distributions $F(\beta,\gamma)$ for the states with $(J,K) = (100,99)$ and $(100,96)$, respectively.

In these high K states, the rotation obviously takes place around the z axis which is perpendicular to the molecular plane and the angular momentum precesses about this axis. In the two panels of Fig. 5.10, the precessing angular momentum forms angles of roughly 9° and 18°, respectively, with the z axis. The angles associated with these precessions compare well with the classical interpretation of the the orientation of the rotation axis relative to the z axis defined as (see, for example, [194, 208])

$$\theta_k = \arccos\left(\frac{\pm K}{\sqrt{J(J+1)}}\right). \tag{5.11}$$

For $K = 99$ and $K = 96$ ($J = 100$) values of θ_k = 10° and 17° are obtained, respectively.

5.4 Discussion

A 'hybrid' line list for $^{32}S^{16}O_3$ has been generated, consisting of 3 414 rotational transitions in the vibrational ground state. The transitions take place between rotational states with J-values in the range from 0 through 85. The frequencies of the 3 414 rotational transitions are obtained using the 'Watsonian' expression of Eq. (5.3) with the values of the $^{32}S^{16}O_3$ ground state rotational constants obtained by Maki *et al.* [104] by least-squares fits to their observed infrared frequencies. The line list values for the intensities of the rotational transitions are purely theoretical, computed with the TROVE program [131,133,165,237] from *ab initio* potential energy and dipole moment surfaces reported previously [170]. The data has been provided in a form suitable for inclusion in the HITRAN database [75] since at the present time of writing, the only SO_3 information available there comprises 25 experimentally measured line frequencies for rotational transitions of $^{32}S^{16}O_3$.

Furthermore, a theoretical investigation has been made into the formation of near-degenerate rotation-vibration energy clusters in the vibrational ground state of $^{32}S^{16}O_3$ at high rotational excitation.

The study shows, for the first time, theoretical predictions of such cluster states

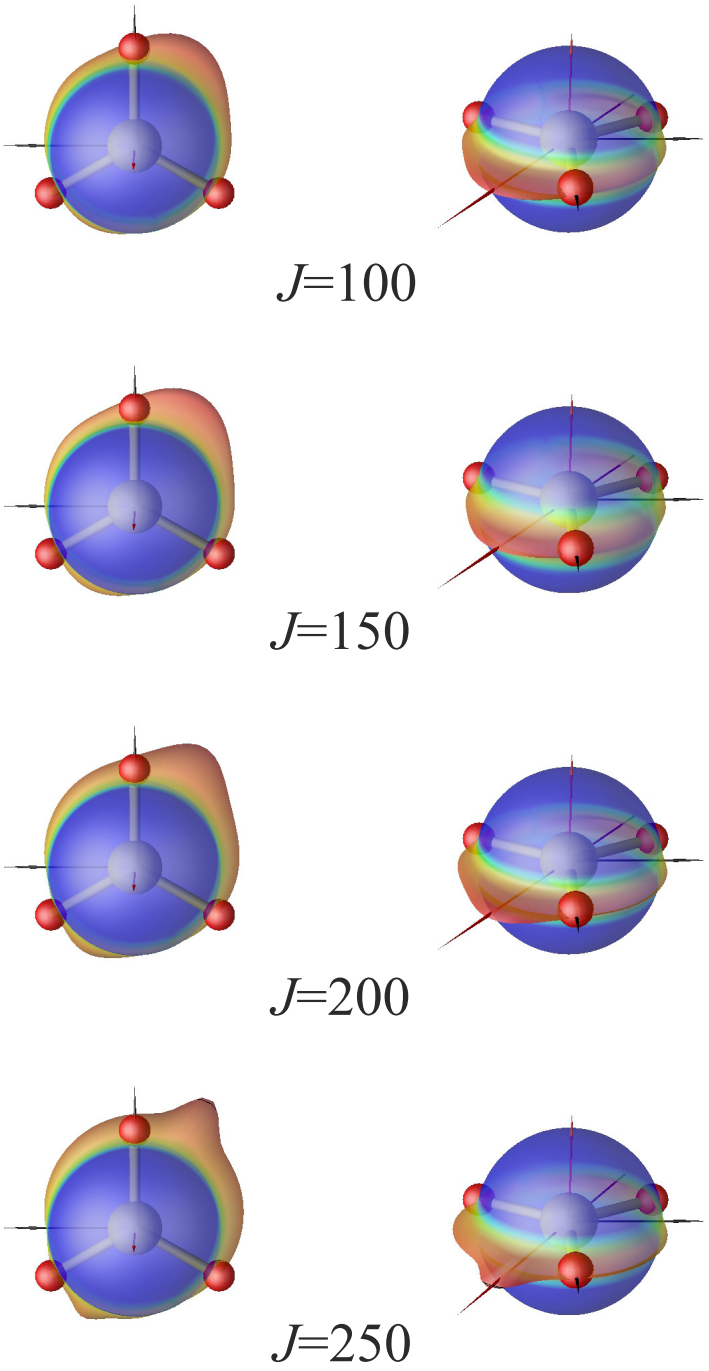

Figure 5.9: Probability distribution functions $F(\beta,\gamma)$ obtained for PCS states with $J = 100$, 150, 200 and 250 in the vibrational ground state of $^{32}S^{16}O_3$. (Multimedia view)

for an XY_3 molecule that is planar at equilibrium. Using variational nuclear motion calculations, the formation of six-fold near-degenerate energy-level clusters in the vibrational ground state of $^{32}S^{16}O_3$ is predicted in the J range between 100 and 250. The structure of the rotational energy clusters is similar to that previously described for

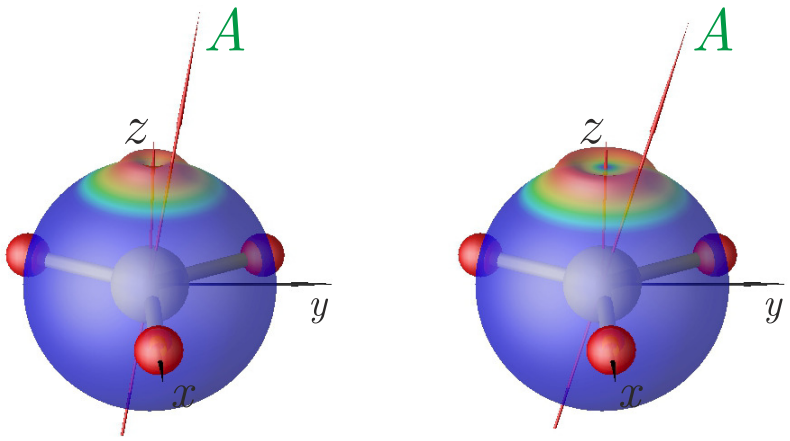

Figure 5.10: Probability distribution functions $F(\beta,\gamma)$ for the rotational states withs $(J,K) = (100,99)$ and $(100,96)$, respectively, in the vibrational ground state of $^{32}S^{16}O_3$.

the pyramidal XH_3 molecules. However the cluster-formation mechanism is different; whereas in the pyramidal molecules the rotation in the cluster states can be thought of as taking place around so-called stabilisation axes roughly coinciding with the X–H bonds, for SO_3 each stabilisation axis is perpendicular to one of the S–O bonds, see Fig. 5.8. In the cluster states at $J \geq 100$, the bond perpendicular to the stable rotational axis is significantly elongated by centrifugal forces.

As the two other bonds are not orthogonal to the axis of rotation, they are less elongated, but the angle α_1 between them is decreased by the centrifugal distortion (Fig. 5.8).

Since the rotational clustering of heavy molecules, such as CF_4 and SF_6, has been successfully studied experimentally, it is hoped that the results of this study will encourage experimentalists to confirm the existence of the 6-fold stabilisation rotations for SO_3 as well. Apart from the conventional spectroscopy at high T, another possibility to reach high rotational excitation experimentally is the optical centrifuge technique [231].

Another reason to consider the rotational clustering SO_3 as unique is that the nuclear statistics permits only the A_1 states, at least in case of the $S^{16}O_3$ isotopologue. This leads to the 6-fold rotational "clustering" represented by a single energy level accompanied by three "dark" levels. It was argued for other clustering systems [238] that the PCS states can be prepared as a long-lived ensemble of eigenstates and thus exper-

imentally observable. Obviously this cannot happen for $S^{16}O_3$, since the components of such an ensemble cannot physically exist. One can speculate about intriguing applications with the optical centrifuge as a sieving device; only isotopologues with the non-zero nuclear spin of oxygen can be driven into the long-lived PCS states.

Chapter 6

The Refinement of the Potential Energy Surface and Computation of a Hot Line List for $^{32}S^{16}O_3$

The computation of a room-temperature line list for $^{32}S^{16}O_3$ provides a starting point for a hot line list calculation. Analyses of the results in Chapter 4 show that while simulated intensities at room-temperature show good reproduction of the 'experimentally derived' absolute intensities, line positions based on a pure *ab initio* calculation are not spectroscopically accurate. The advantage of the room-temperature line list calculation lies in the fact that these discrepancies can be observed for a lower computational expense, allowing for modifications to be made to minimise the disagreements before computing the final hot line list.

To be able to simulate spectra at higher temperatures, it is generally necessary to increase the basis set upon which rovibrational calculations are performed. As well as increasing the accuracy of band positions, the larger basis set size can account for higher energy bands which become more populated as the temperature increases. The consequence of this, however, is that the line list computation becomes a much more intimidating task; the matrices which require diagonalization are larger, and their wavefunction solutions are composed of linear combinations of a greater number of basis functions. This means the demand for computational resources is greater, with more time, memory and storage space required for all calculations.

In this chapter the hot line list, UYT2 is presented. Discussed are the refinement

procedure to the potential energy surface, the increased basis set, and the methods employed to overcome the computational difficulties in order to ensure complete construction of the line list.

6.1 Refining the Potential Energy Surface

The first stage in the computation of the hot line list focuses on the refinement of the *ab initio* PES. This involves performing a least-squares fit of *ab initio* rovibrational energies to experimentally obtained data. The method follows that described in Refs. [130, 239–241], and uses the program TROVE [131]. The essence of the procedure is to introduce a linear correction term to the *ab initio* Hamiltonian in order to minimise a functional that quantifies the differences between the *ab initio* and experimental values compared.

In Chapter 4 the analytical form of the *ab initio* PES, V_0, is described in Equation 4.1, expanded in terms of a set of internal coordinates (Equations 4.2). The refinement procedure assumes that this is a reasonable approximation to the 'true' PES; a correction, ΔV, is introduced to this PES, also expanded in terms of the internal coordinates described above, i.e. $H = H_0 + \Delta V$, where the unperturbed Hamiltonian $H_0 = T + V_0$ contains the kinetic energy, T, and the unrefined PES, V_0.

The refined potential coefficients introduced by the 'perturbation' are given by

$$F'_{jkl\ldots} = F_{jkl\ldots} + \Delta F_{jkl\ldots}, \tag{6.1}$$

where the $F_{jkl\ldots}$ are in accordance with the expansion coefficients given by Equation 4.1. The determination of the fitting parameter $\Delta F_{jkl\ldots}$ is achieved from a least squares fitting algorithm based on the derivatives of energies with respect to $\Delta F_{jkl\ldots}$ computed via the Hellman-Feynman theorem [242]. The process is iterative; starting with a value of zero for all $\Delta F_{jkl\ldots}$, optimal values for these perturbation expansion coefficients can be obtained by allowing them to vary until adequate agreement is found between the energies obtained from the calculation and their experimental counterparts.

In practice, not all of the $\Delta F_{jkl\ldots}$ require varying in order to achieve a minimization between observed and calculated energies; each of the expansion coefficients in Eq. 6.1 is associated with specific vibrational coordinates that are included in the expansion of

the PES, and therefore with the rovibrational energies described by basis functions in these coordinates. In addition to this, the result of the variation of one parameter may or may not be enhanced by the variation of another. As such, the minimisation scheme is largely a trial and error process in which multiple combinations of varying parameters must be tested before arriving at the optimal result.

It should be noted that the refined PES is 'effective' in that it is dependent on the basis set used, i.e. the results presented here may only be obtained from the implementation of the refined PES into TROVE in conjunction with the basis set and operator truncation schemes (see Chapters 2 and 4) that are used in the refinement procedure. The truncation parameters employed in the calculation of UYT2 are specified in Section 6.2.

6.1.1 Experimental Data Used in the Refinement

The experimental data for the energies is taken from the work of Maki *et al.* [99–105] which has been discussed in Chapter 4. Despite being the only reputable data set available at the time that the current work was carried out, the advantage is that the it can be assumed to be more self-consistent compared to spectroscopic data that may be found in collaborative databases. The majority of the data of Maki *et al.* provides upper and lower energy states labelled by their vibrational and rotational quantum numbers, based on a fitting procedure to an effective Hamiltonian.

The refinement procedure includes rotationally excited states which facilitates improvement from the perspective of the entire band. Since the energy levels used in the Hamiltonian fits of Maki *et al.* are predominantly rotationally excited levels, then the majority of the levels used in the refinement procedure are also for $J > 0$, therefore minimizing any ambiguity in the band positions. A total of 119 energy levels for $J \leq 5$ were chosen from this set based on their reliability at reproducing the observed transitions, with the condition that they are physically accessible states with A' or A'' symmetry; any published values of experimentally-derived purely vibrational terms (i.e. band centres) that are infra-red inaccessible were not included. Table 6.1 lists all the $J = 5$ levels used in the refinement process, comparing with their final computed counterparts.

Table 6.2 shows the effect of the final potential refinement on the bands used in the refining procedure. The root mean square (RMS) differences are calculated by match-

Table 6.1: Obs. - Calc. residuals for the $J = 5$ energy levels used in the refinement procedure. All values are in cm^{-1}. The corresponding values for $J < 5$ are included in the appendix.

State	K	Obs. [104]	UYT2	Obs. - Calc.
ν_0	3	8.885	8.886	-0.001
ν_2	3	506.367	506.360	0.008
	0	507.900	507.893	0.007
	5	535.323	535.312	0.011
$\nu_4^{(l_4=1)}$	4	538.471	538.490	-0.020
	2	539.561	539.560	0.001
	1	540.677	540.685	-0.008
$2\nu_2$	3	1002.357	1002.411	-0.054
$\nu_2 + \nu_4^{(l_4=1)}$	5	1033.102	1033.058	0.044
	4	1036.241	1036.226	0.016
	2	1037.252	1037.219	0.033
	1	1038.243	1038.220	0.023
	5	1068.278	1068.303	-0.025
$2\nu_4^{(l_4=0)}$	3	1068.461	1068.456	0.005
$2\nu_4^{(l_4=2)}$	2	1071.024	1071.031	-0.007
	5	1398.427	1398.437	-0.010
$\nu_3^{(l_3=1)}$	2	1401.580	1401.581	-0.001
	1	1401.599	1401.591	0.009
	5	1529.365	1529.362	0.003
$2\nu_2 + \nu_4^{(l_4=1)}$	4	1532.498	1532.520	-0.022
	2	1533.442	1533.448	-0.006
$\nu_1 + \nu_2$	3	1573.870	1573.856	0.014
	0	1575.400	1575.387	0.013
$3\nu_4^{(l_4=1)}$	4	1597.408	1597.410	-0.002
$\nu_1 + \nu_4^{(l_4=1)}$	5	1601.162	1601.150	0.012
	4	1604.308	1604.322	-0.014
	2	1605.430	1605.426	0.004
	1	1606.574	1606.577	-0.002
	5	1923.797	1923.808	-0.011
$^\star(\nu_3 + \nu_4)^{(L=2)}$	4	1925.310	1925.318	-0.008
	0	1927.488	1927.422	0.066
	1	1927.982	1927.988	-0.006
	5	2782.262	2782.227	0.035
$2\nu_3^{(l_3=2)}$	4	2786.812	2786.837	-0.025
	2	2786.901	2786.888	0.014
	1	2788.419	2788.425	-0.006
$3\nu_3^{(l_3=1)}$	5	4143.316	4143.246	0.070
	2	4146.379	4146.329	0.050

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

ing all experimental lines within each band with calculated values via their quantum number assignments.

Table 6.2: Comparison of calculated (TROVE - UYT and UYT2) and experimental [104] bands of $^{32}S^{16}O_3$.

Band	Band Centre (cm^{-1})		
	Obvs.	UYT_{RMS}	$UYT2_{RMS}$
$2\nu_2$ - ν_2	497.45	0.73	0.09
ν_2 - ν_0	497.57	0.77	0.05
$\nu_2 + \nu_4$ - ν_4	497.81	0.82	0.03
$2\nu_4^{(l_4=0)}$ - ν_4	529.72	1.33	0.30
ν_4 - ν_0	530.09	1.41	0.09
$\nu_2 + \nu_4$ - ν_2	530.33	0.22	0.08
$2\nu_4^{(l_4=2)}$ - ν_4	530.36	1.54	0.37
ν_1 - ν_4	534.83	0.47	0.20
ν_3 - ν_0	1391.52	4.06	0.09
$2\nu_2 + \nu_4$ - ν_0	1525.61	0.19	0.08
$^{\S}\nu_2 + 2\nu_4^{(l_4=0)}$ - ν_0	1557.88	2.39	1.17
$^{\S}\nu_2 + 2\nu_4^{(l_4=2)}$ - ν_0	1558.52	2.12	0.64
$^{\S}\nu_1 + \nu_2$ - ν_0	1560.6	1.14	1.28
$^{\S}3\nu_4^{l_4=1}$ - ν_0	1589.81	6.73	4.00
$^{\S}\nu_1 + \nu_4^{(l_4=1)}$ - ν_0	1593.69	3.32	3.57
$^{\star\S}(\nu_3 + \nu_4)^{(L=2)}$ - ν_0	1917.68	5.34	0.65
$2\nu_3^{(l_3=2)} - \nu_0$	2777.87	7.53	0.20
$^{\S}3\nu_3^{(l_3=1)}$ - ν_0	4136.39	–	0.08

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].
§These bands are refined by using energy levels obtained via a combination difference method.

The bands studied in Ref. [105] label transitions by rotational and vibrational quantum numbers, but do not list upper and lower energy states. To obtain states for these bands, combination difference methods were used using knowledge of the transition frequency and pure rotational terms values of the lower state (vibrational ground state). The bands which have been refined using this method are highlighted in Table 6.2.

It should be noted that in matching line positions between the experimental and computed values, a number of data points were not included in the calculation of the RMS differences. These data points coincide with transitions which were not included in the Hamiltonian fitting algorithm of Maki *et al.*

The RMS differences calculated are slightly degraded for a few bands as a result of the refinement, namely the ν_1 + ν_2 - ν_0 and ν_1 + $\nu_4^{(l_4=1)}$ - ν_0 bands. The experimental energy levels used to refine these two bands are obtained from the combination difference method mentioned above, for a number of low-lying rovibrational states. However, for more rotationally excited levels within these bands, the quantum number labelling for experimental transitions is slightly dubious; a number of transitions are labelled with identical upper and lower rotational quantum numbers. Transitions with identical quantum numbers are excluded from the RMS difference calculations, although this alone does not account for the diminished RMS values. However, the presence of duplicate quantum number assignments may indicate incorrect labelling within the rest of the band, and render the RMS values dubious. Without specification of the upper and lower experimental energy wavenumbers involved in the transitions it is difficult to ascertain the source of the discrepancy. The experimental data is included in the supplementary material of Ref. [105].

Table 6.3 compares all the published purely vibrational terms (J = 0) with those calculated with TROVE before and after refinement. The are a number of notable discrepancies introduced by the refinement procedure and in some cases deteriorations from the pre-refined values (e.g $2\nu_2$). The inclusion of terms which are not IR accessible tend to cause major deviations in the PES fit. Since the majority of these terms are physically inaccessible (and therefore only inferred) it is more pertinent to assess the quality of the refinement via Table 6.2,with the exception of $3\nu_2$, for which there is no experimental band data available beyond the quoted vibrational term value.

6.2 Calculation using TROVE

The method of calculating rovibrational energies and transition intensities for the hot line list is, in essence, identical to that outlined in Section 4.1 of Chapter 4, by use of the TROVE program [131]. The main differences lie in both the use of the refined CCSD(T)-F12b/aug-cc-pVTZ-F12 potential energy surface (described above) and an increase in the size of the basis set used in the variational calculations. The same *ab initio* Dipole Moment Surface (DMS) used in the UYT calculation is also used for UYT2 (see Chapter 4).

Various truncation parameters are described in Section 4.1 of Chapter 4 with re-

Table 6.3: Comparisons of purely vibrational (J = 0) terms for $^{32}S^{16}O_3$, between experimentally published values, and pre- and post-refinement of the *ab initio* PES using TROVE (UYT and UYT2, respectively). All values are given in cm^{-1}.

	Obs. [104]	UYT	UYT2
ν_2	497.57	498.48	497.56
ν_4	530.09	528.59	530.09
$2\nu_2$	995.02	995.35	993.67
$\nu_2 + \nu_4^{(l_4=1)}$	1027.90	1027.35	1027.33
$2\nu_4^{(l_4=0)}$	1059.81	1056.50	1059.48
$2\nu_4^{(l_4=2)}$	1060.45	1057.38	1060.45
ν_1	1064.92	1065.75	1066.49
ν_3	1391.52	1387.45	1391.51
$3\nu_2$	1492.35	1490.76	1488.47
$2\nu_2 + \nu_4^{(l_4=1)}$	1525.61	1524.48	1524.20
$\nu_2 + 2\nu_4^{(l_4=0)}$	1557.88	1555.59	1557.50
$\nu_2 + 2\nu_4^{(l_4=2)}$	1558.52	1556.45	1558.46
$\nu_1 + \nu_2$	1560.60	1565.33	1565.07
$3\nu_4^{(l_4=1)}$	1589.81	1586.46	1588.97
$3\nu_4^{(l_4=3)}$	1591.10	1586.43	1591.06
$\nu_1 + \nu_4^{(l_4=1)}$	1593.69	1593.36	1595.92
$\nu_2 + \nu_3^{(l_3=1)}$	1884.57	1881.53	1884.29
$^\star(\nu_3 + \nu_4)^{(L=2)}$	1917.68	1912.24	1917.68
$^\star(\nu_3 + \nu_4)^{(L=0)}$	1918.23	1914.56	1919.63
$2\nu_3^{(l_3=0)}$	2766.40	2759.12	2766.38
$2\nu_3^{(l_3=2)}$	2777.87	2770.29	2777.86
$3\nu_3^{(l_3=1)}$	4136.39	4126.78	4136.33

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

spect to the basis set used in computing the room-temperature line list, namely the kinetic/potential expansion and polyad number truncation schemes. For the computation of UYT2 the parameters N_{kin}, P_{max} were set to 6 and 18 (instead of 4 and 12 used for UYT), respectively. The larger basis set allows for greater convergence at higher energies which become important as temperature increases.

The room-temperature line list is also expanded by computing all physically allowed rovibrational wavefunctions of A_1' and A_1'' symmetry corresponding to energies satisfying $E < 10\ 000\text{cm}^{-1}$. The final line list consists of all transitions between 0 $< \nu \leq 5000$ cm^{-1}, satisfying the conditions $E'_{max} \leq 9000$ cm^{-1}, $E''_{max} \leq 4000$ cm^{-1}. Rotational excitations for all $J \leq 130$ are considered. The expansion described here is

necessary, along with the basis set increase, in order to guarantee complete coverage and accuracy when simulating high temperatures. However, as discussed previously, the need for greater accuracy comes at a computational cost. As such, the computation of UYT2 is significantly more demanding than for the room-temperature list.

Construction of the line list took the best part of 2 years, owing to the increased computation time in the calculation of wavefunction files and transition strength calculations, in addition to the queuing systems imposed on job submissions, and the resubmission of failed jobs.

6.2.1 Increasing Computational Demands

6.2.1.1 Matrix Construction and Diagonalisation

In terms of memory, the diagonalisation of the constructed Hamiltonian matrices is the most computationally expensive part of the line list calculation. For each J, a matrix is built as described in Refs. [131, 165] (and Chapter 2) and then stored in the memory for diagonalisation, using an appropriate eigensolver routine. Since only the A_1' and A_1'' symmetry species are required, it is possible to separate each J matrix further into these symmetry blocks to be dealt with separately.

The dimension of each matrix is indicative of the memory requirements needed to store it, and roughly scales as $N_{J=0} \times (2J+1)$, where $N_{J=0}$ is the dimension of the purely vibrational matrix. For UYT2, the combined dimension $N_{J=0}$ of both A_1' and A_1'' symmetries is 2692; to put this into perspective, the same value in the UYT line list calculations is $N_{J=0} = 679$. The size of the largest matrix considered in the room-temperature calculations (for $J = 85$) is $N_{J=85} = 111296$, which is a value already surpassed by $J = 21$ for the hot line list (for UYT2, the value of $N_{J=85}$ is 454488). It became quickly apparent that the diagonalisation techniques previously employed to determine the UYT wavefunctions would not be sufficient in the case of UYT2; full diagonalisation scales roughly as N^3 on CPU time, and N^2 on memory requirement.

As well as a large memory requirement, both increased CPU resources and adequate storage space for the produced wavefunction files are necessary for the hot line list calculation. The computational resources described in Chapter 4 used in the UYT calculations could not provide either of these in suitable amounts, and therefore calculations were performed elsewhere on both the Darwin and COSMOS HPC facilities in

Cambridge.

Each of the computing nodes on the Darwin cluster provide 16 CPUs across two 2.60 GHz 8-core Intel Sandy Bridge E5-2670 processors, and a maximum of 64 Gb of RAM. The advantage of moving eigenfunction calculations to the Darwin cluster are that an entire node can be dedicated to one calculation, spread across the 16 CPUs, which is double the highest amount that was available to a calculation for the room-temperature line list (see Chapter 4). Since multiple nodes can be accessed by a single user at any time, multiple computations can be carried out simultaneously.

Diagonalisations for a number of low J matrices were possible using the same approach as the room-temperature line list, employing the LAPACK DSYEV eigensolver [186], optimised for OpenMP parallelisation across multiple CPUs (in this case 16). For $J \leq 32$ this method was used, before constraints due to the memory (and to a lesser extent, time) limitations made this no longer possible.

For $32 < J \leq 90$ a shared memory approach was used with an MPI-optimised version of the eigensolver, PDSYEVD, which allowed diagonalisation across multiple Darwin/COSMOS nodes in order to make use of their collective memory. In order to diagonalise the matrix within the 36 hour wall clock limit, it was necessary to perform this method in three steps. First, for a given J and symmetry species Γ, the Hamiltonian matrix is constructed and saved to disk. Secondly, the matrix is then read and diagonalised in the PDSYEVD routine across a specified number of nodes capable of containing the matrix within their shared memory. This produces a set of checkpoint files which must be finally read in again by another process and saved in the TROVE eigenfunction format.

For $J < 90$ yet another approach was developed for use on the COSMOS shared memory machine. This method employs the PLASMA DSYTRDX routine and, unlike the above method, can construct, diagonalise and store wavefunctions to disk in one single process. However, the memory and time requirements did not fall within the standard user restrictions of the cluster, and special access to a larger amount of compute nodes and overall wall clock time were required. For $J = 130$ ($\Gamma = A_1'$) a total of 52 hours of real time was taken to construct and diagonalize the Hamiltonian matrix across 416 CPUs, and utilising 3140 Gb of RAM. The computational demands of such a process severely limited the number of simultaneously running jobs allowed at

a given time.

Figure 6.1 shows the linear scaling of matrix dimension with J, as well as the number of eigenvalues under the 10 000 cm^{-1} threshold, which does not scale linearly with J. The size of the produced wavefunction files follow the same trend, and peak between $60 < J < 70$ at around 300 Gb for both A_1' and A_1'' symmetries combined. It was impossible to compute and store the required wavefunction files all at once as their large size put considerable pressure on the storage constraints of the systems in use. As such, the process was staggered, as well as performed on multiple systems.

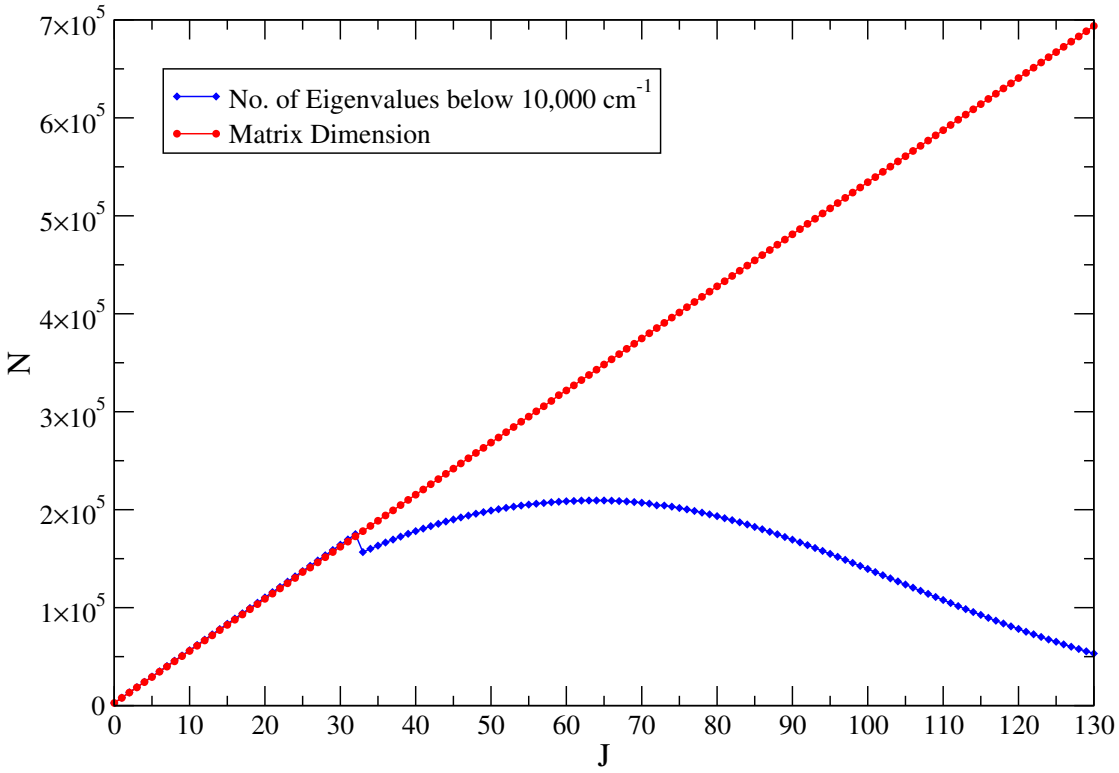

Figure 6.1: Size of Hamiltonian matrix (A' and A'' symmetries combined) and number of eigenvalues below 10 000 cm^{-1} as a function of increasing J. The discontinuity at J = 32 is a result of the change in diagonalization procedure; before this point all eigenfunctions of the Hamiltonian are saved, whereas beyond this point only those below 10 000 cm^{-1} are kept.

6.2.1.2 Calculation of Transition Strengths

Where memory is the major requirement in the matrix diagonalisation stage of the line list calculation, computation time is the largest component of the transition strength phase. There are a number of methods for reducing the computational demands. Firstly, calculations of the line strength $S(f \leftarrow i)$ only take into consideration the basis functions of the final state wavefunction who's coefficients are greater than 10^{-4}. In addi-

tion to this a threshold value for the Einstein A coefficient of 10^{-74} s^{-1} dictates which transitions are kept. However, the number of line strength calculations to be performed is still very large.

Using TROVE, it is possible to parallelise this process by separating calculations between each $J \leftrightarrow J+1$ pair and run them on different nodes simultaneously. However, due to the huge density of states for each J in the UYT2 calculation, it was necessary to separate these further into windows of lower energy states, in order to ensure all transitions for a given $J \leftrightarrow J+1$ pair can be computed within a wall clock limit. For example, for one $J \leftrightarrow J+1$ pair, one process may compute all transitions between $0 < E'' \leq 1000$ cm^{-1} and $0 < E' \leq 10\,000$ cm^{-1}, while another computes all transitions between $1000 < E'' \leq 2000$ cm^{-1} and $0 < E' \leq 10\,000$ cm^{-1}, etc.

However, even with these parallelisation techniques employed across multiple Darwin CPUs, the calculations were still subject to a long, arduous process; the multiple submission of parallelised jobs all demanding maximum wall clock time limits meant that the job submissions were subject to extended queue times based on a fair-usage policy of the cluster. In addition to this, particularly in calculations involving high densities of states, many jobs would be terminated before completion due to reaching the imposed time limit.

In order to help alleviate this an adapted version of TROVE was used, optimised for performing calculations on graphical processing units (GPUs). The use of this implementation, known as GAIN [243], allowed for the computation of transition strengths for the more computationally demanding parts of the calculations. These were performed on the Emerald GPU cluster, based at CFI in Southampton. In general, the calculation of transition strengths across multiple GPUs was much faster than the Darwin CPUs. For example, there are a total of 349 481 979 transitions for J'' = 35, which took a total of 17 338 CPU hours to compute on the Darwin nodes, compared to 2053 GPU hours on the Emerald nodes for 346 620 894 transitions for J'' = 59. These GPU calculations were carried out for $J \leftrightarrow J+1$ pairs containing a large number of states, while the Darwin CPUs were reserved for the less computationally demanding sections.

Despite these measures, submitted jobs were still prone to termination due to various hardware issues and wall clock limits, as mentioned above. It was necessary to

either resubmit jobs that had been terminated, or continue from where they left off, which required the use of various techniques to ensure that duplicate lines were not calculated, including the use of shell scripts and, in many cases, manual checking.

Figure 6.2 shows the number of transitions between J pairs as a function of J''. The number of transitions peaks in the region $35 < J < 65$, where the most computationally expensive part of the calculations resided.

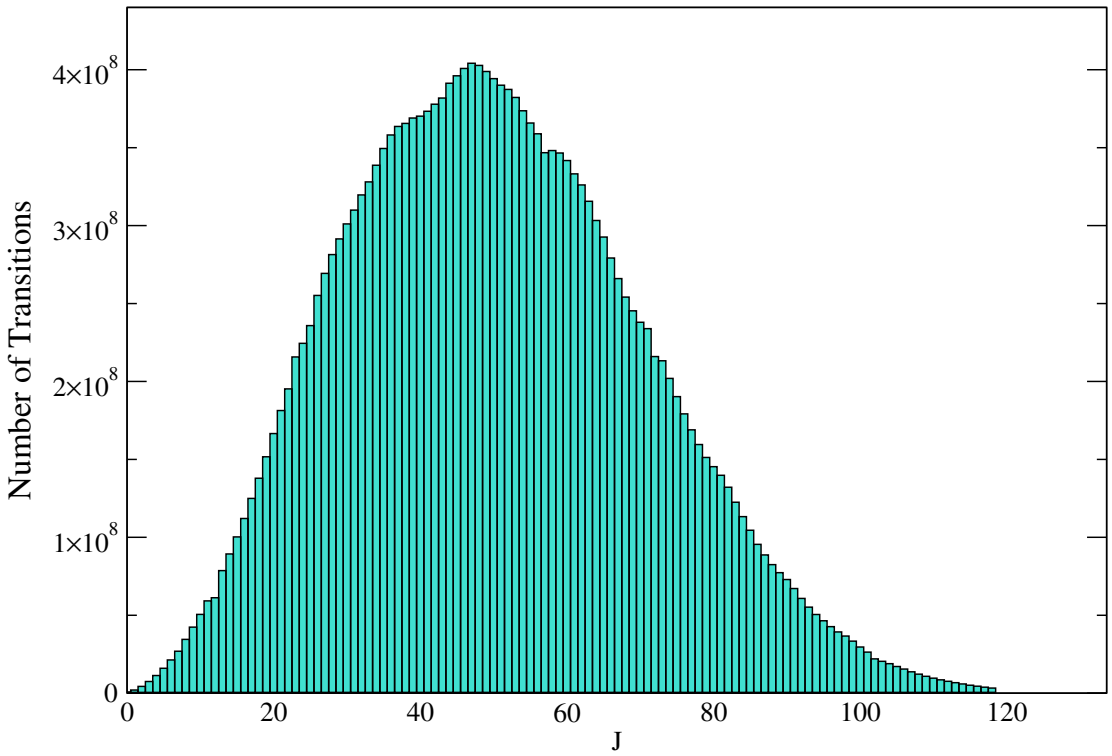

Figure 6.2: Plot of the number of transitions calculated from energy levels with rovibrational quantum number J.

The total number of transitions calculated for UYT2 satisfying $E'' \leq 4000$ cm^{-1} and within the region $0 < \nu \leq 5000$ cm^{-1} is 21 billion, an amount which is two orders of magnitude larger than UYT. Despite the described measures taken to spread the computational demand across all available resources, the lower energy cut-off of 4000 cm^{-1} is a compromise; the original aim was to include transitions from higher energy states. Unfortunately the ever increasing density of states beyond 4000 cm^{-1} made these ambitions impossible within a reasonable amount of time. Throughout the project much time was spent on attempting to estimate the computational time required to calculate transitions from such high states, and whether it was feasible given the resources available. Since it is difficult to make such estimations *a priori* the process

Table 6.4: Extract from the UTY2 state file for SO_3; quantum numbers are specified in Table 6.5.

i	E	g	J	Γ_{Total}	K	Γ_{Rot}	n_1	n_2	n_3	n_4	n_5	n_6	v_1	v_2	v_3	L_3	v_4	L_4	Γ_{Vib}
1	0.0000	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
2	993.6780	1	0	1	0	1	0	0	0	0	0	2	0	2	0	0	0	0	1
3	1059.4770	1	0	1	0	1	0	0	0	0	2	0	0	0	0	0	2	0	1
4	1066.4970	1	0	1	0	1	1	0	0	0	0	0	1	0	0	0	0	0	1
5	1591.0349	1	0	1	0	1	0	0	0	0	3	0	0	0	0	0	3	1	1
6	1919.6346	1	0	1	0	1	1	0	0	0	1	0	0	0	1	1	1	1	1
7	1981.9944	1	0	1	0	1	0	0	0	0	0	4	0	4	0	0	0	0	1
8	2054.0505	1	0	1	0	1	0	0	0	0	2	2	0	2	0	0	2	0	1
9	2061.9334	1	0	1	0	1	1	0	0	0	0	2	1	2	0	0	0	0	1
10	2117.4659	1	0	1	0	1	0	0	0	0	4	0	0	0	0	0	4	0	1
11	2124.4973	1	0	1	0	1	1	0	0	0	2	0	1	0	0	0	2	0	1
12	2129.3331	1	0	1	0	1	0	1	1	0	0	0	2	0	0	0	0	0	1
13	2444.1614	1	0	1	0	1	1	0	0	0	2	0	0	0	1	1	2	2	1
14	2586.0493	1	0	1	0	1	0	0	0	0	3	2	0	2	0	0	3	1	1
15	2648.2382	1	0	1	0	1	0	0	0	0	5	0	0	0	0	0	5	1	1
16	2655.7551	1	0	1	0	1	1	0	0	0	3	0	1	0	0	0	3	1	1
17	2766.3812	1	0	1	0	1	0	2	0	0	0	0	0	0	2	0	0	0	1
18	2904.3481	1	0	1	0	1	1	0	0	0	1	2	0	2	1	1	1	1	1

was largely based on trial and error, which resulted in the submission of thousands of jobs all subject to long queue times, many of which resulted in termination before their completion.

Ultimately, full coverage of the final line list has been sacrificed by the reduction of lower energy states used, in order for the computation to have been carried out within the time constraints of the project. Section 6.4 discusses the extent of this coverage/resource trade-off and the effect it has upon the simulation of rovibrational spectra.

6.3 Overview of the UYT2 Line List

The UYT2 line list is presented in the ExoMol format [175] described in Chapter 2, i.e. via a 'states', containing a list of energy levels, and a 'transition' file which lists upper and lower energies via their index from the 'states' file, and associated Einstein A coefficient. Tables 6.4 and 6.6 give portions of these files.

The energy levels listed in the 'states' file are labelled with the quantum numbers summarised in Table 6.5.

Table 6.5: Quantum numbers used in labelling energy states.

Quantum Number	Description
i	Energy index
E	Energy value (cm^{-1})
g	Total degeneracy of the state
J	Rotational angular momentum quantum number
Γ_{Total}	Total symmetry in $\mathscr{D}_{3h}$(M)
K	Projection of angular momentum on to the z-axis
Γ_{Rot}	Rotational symmetry in $\mathscr{D}_{3h}$(M)
v_1, v_2, v_3, v_4	Normal mode vibrational quantum numbers
L_3, L_4	v_3 and v_4 projections of the angular momenta
Γ_{Vib}	Vibrational symmetry in $\mathscr{D}_{3h}$(M)
n_1, n_2, n_3, n_4, n_5, n_6	Local mode vibrational quantum numbers

Table 6.6: Extract from the UTY2 transitions file for SO_3.

f	i	A
237007	249581	1.1253e-17
158430	148459	2.8358e-17
549592	568676	1.3725e-16
120670	112002	1.4546e-16
2080392	2117071	9.0696e-18
289088	302965	1.4938e-16
393104	377035	1.5764e-16
43637	49289	2.1375e-16
587986	607961	2.0370e-16
587868	647986	4.2068e-18
2007259	2043487	5.2490e-18
627725	648113	3.0673e-16

All quantum numbers are provided by TROVE, with the exception of the normal mode vibrational quantum numbers, v_1, v_2, v_3 and v_4, and their angular momentum projections $L_3 = |l_3|$ and $L_4 = |l_4|$, which are obtained from the local mode quantum numbers via the following rules:

$$v_1 + v_3 = n_1 + n_2 + n_3, \qquad v_2 + v_4 = n_4 + n_5 + n_6, \tag{6.2}$$

and

$$l_3 = -n_3, -n_3 + 2, ..., n_3 - 2, n_3, \qquad l_4 = -n_4, -n_4 + 2, ..., n_4 - 2, n_4. \tag{6.3}$$

The mapping between these quantum numbers for a particular level also required

knowledge of the energy value and symmetry, since multiple levels may be labelled with identical local mode quantum numbers. In these cases, the symmetric modes ν_1 and ν_2 are associated with the lower energy, and ν_3 and ν_4 with the higher, and that L_3 and L_4 increase proportionally with the energies, and are multiples of 3 in the case of A_1 or A_2 symmetries, or otherwise for E-type symmetry. This mapping is performed by hand at the $J = 0$ stage of the calculation, and then extrapolated to $J > 0$.

Of all the quantum numbers (except the running index), only J, g_{Total} and Γ_{Total} are rigorously defined. The remaining quantum numbers represent the 'largest contribution' from rotational and vibrational components of the wavefunction expansion associated with a given state.

The 'transition' file contains N_i and N_f, the indexes of an initial and final state, respectively, in conjunction with the 'states' file, and the Einstein A coefficient in s^{-1} associated with the transition. The rigorous selection rules for electric dipole transitions, $\Delta J = J' - J'' = 0, \pm 1$ $(J'' + J' \geq 1)$, are obeyed, as are the symmetry selection rules $A_1' \leftrightarrow A_1''$.

6.4 Temperature Dependence and Completeness of the Line List

As described in Chapter 2, in order for a line list to be suitable for modelling spectra at a certain temperature it is necessary for the partition function, Q, to be converged at this temperature. This is equivalent to stating that all energy levels that are significantly populated at the given temperature, T, must be computed. This is the metric upon which the line list completeness is gauged. The analysis of the partition function in this section follows a similar procedure to that of Chapter 3 for $^{32}S^{16}O_2$.

In Chapter 4, for the room-temperature $^{32}S^{16}O_3$ line list, convergence was found for Q at a value of 7908.266 for the absolute temperature T = 298.15 K, by including all rovibrational energies for $J \leq 85$ using the described basis set. However, as the temperature increases, so does the contribution to Q from the higher lying states, therefore the number of energies required increases. This is the reason for the expansion in basis set size and calculation of higher rotationally excited states in UYT2, the 'hot' line list.

Figure 6.3 shows convergence of the partition function as a function of J_{max} for

different absolute temperatures, T.

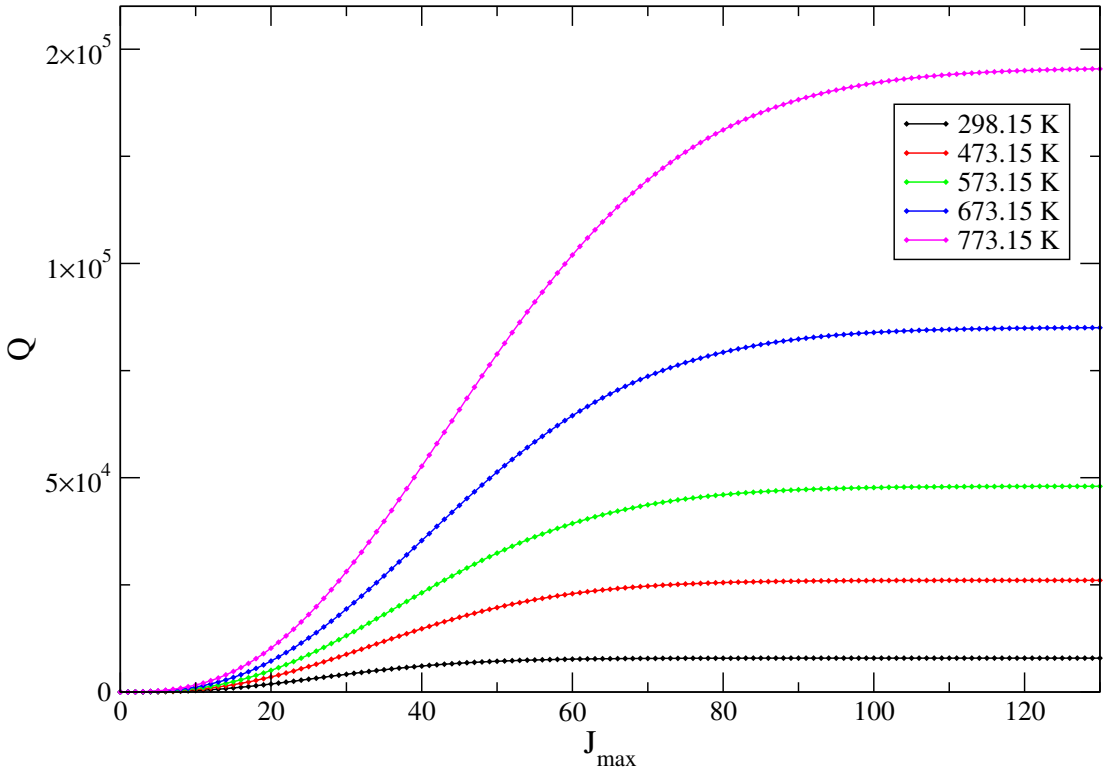

Figure 6.3: Convergence of partition function at different temperatures as a function of J_{max}.

Upon inspection, the value of Q for each value of T is adequately converged at J = 130. Table 6.7 shows the final values of Q obtained for each temperature, quoted alongside their degree of convergence. As can be seen, the value of Q = 7908.906 at T = 298.15 K calculated from UYT2 is in very good agreement with the value of Q = 7908.266 obtained from UYT.

Table 6.7: Values of the partition function, Q, for different temperatures, T. The degree of convergence is specified by $Q_{J_{130}} - Q_{J_{129}}/Q_{J_{130}} \times 100$.

T (K)	Q	Degree of Convergence (%)
298.15	7908.906	6.27×10^{-6}
473.15	26065.642	8.50×10^{-4}
573.15	48007.866	3.62×10^{-3}
673.15	85016.645	9.99×10^{-3}
773.15	145389.574	2.12×10^{-2}

For the purposes of determining the completion of the line list, it is perhaps more appropriate to view the convergence of Q as a function of an energy cut-off, E_{max} (in the same vein as Chapter 3 for the case of $^{32}S^{16}O_2$). This is shown in Figure 6.4.

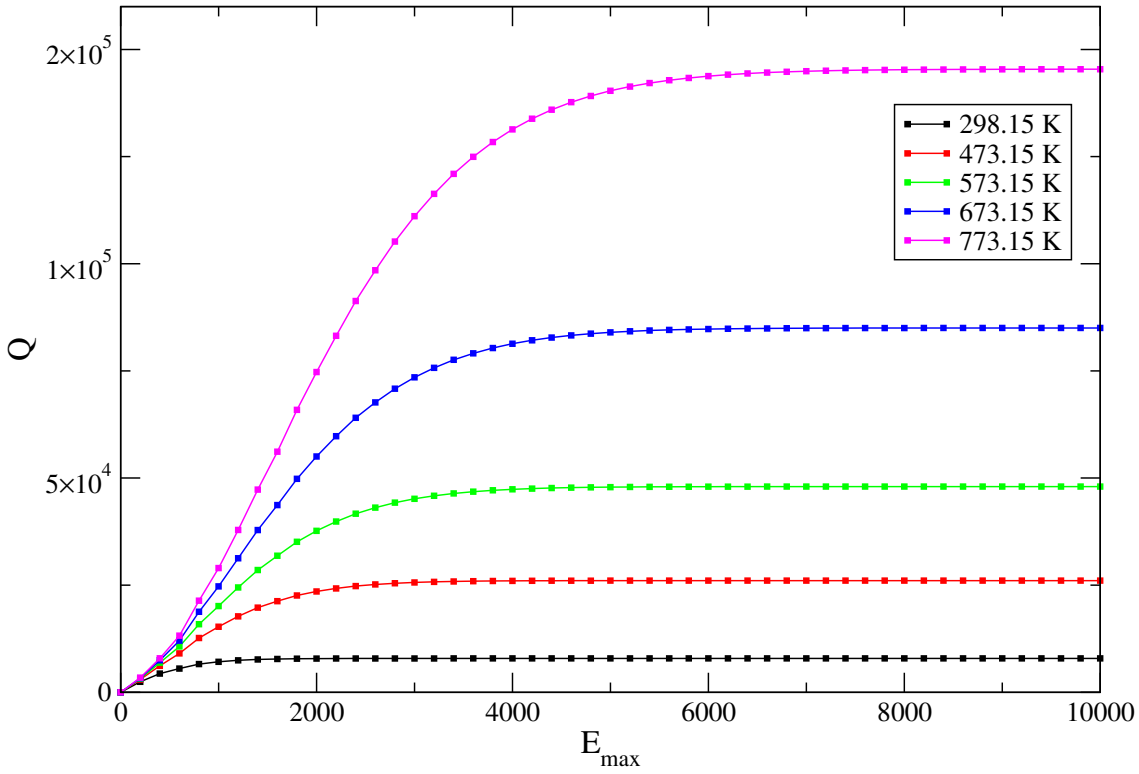

Figure 6.4: Convergence of partition function at different temperatures as a function of E_{max} (cm^{-1}).

UYT2 only contains transitions from levels with $E'' \leq 4000$ cm^{-1}, as a consequence of the computational and time limitations discussed in the Section 6.2. It is clear from Figure 6.4 that imposing this limit will have a non-negligible effect on a spectral simulation at $T = 773.15$ K, in particular; since the partition function is not fully converged at $E_{max} = 4000$ cm^{-1} it is expected that levels with energies above this value will also be populated to some extent. This would be manifest as certain lines being missing from the spectrum, where transitions from levels contributing with some significance to the partition function are not included.

Similarly, the truncation of calculations at $J = 130$ means that a number of potentially contributing energy levels have been completely neglected from the partition sum at $T = 773.15$ K; at $J = 130$ the lowest energy lies around 4000 cm^{-1}. Therefore, as well as transitions from these states not appearing on the simulated spectrum, their absence also means that the partition function obtained here will be slightly lower than the fully converged value.

As in Chapter 3, it is possible to quantify the completeness of the line list by assuming that the value of Q at $J = 130$ is close enough to the 'true' value of the

partition function at the given temperature. Figure 6.5 shows the ratio of the value of the partition function at the 4000 cm^{-1} cut-off and the assumed total partition function, Q_{Total}.

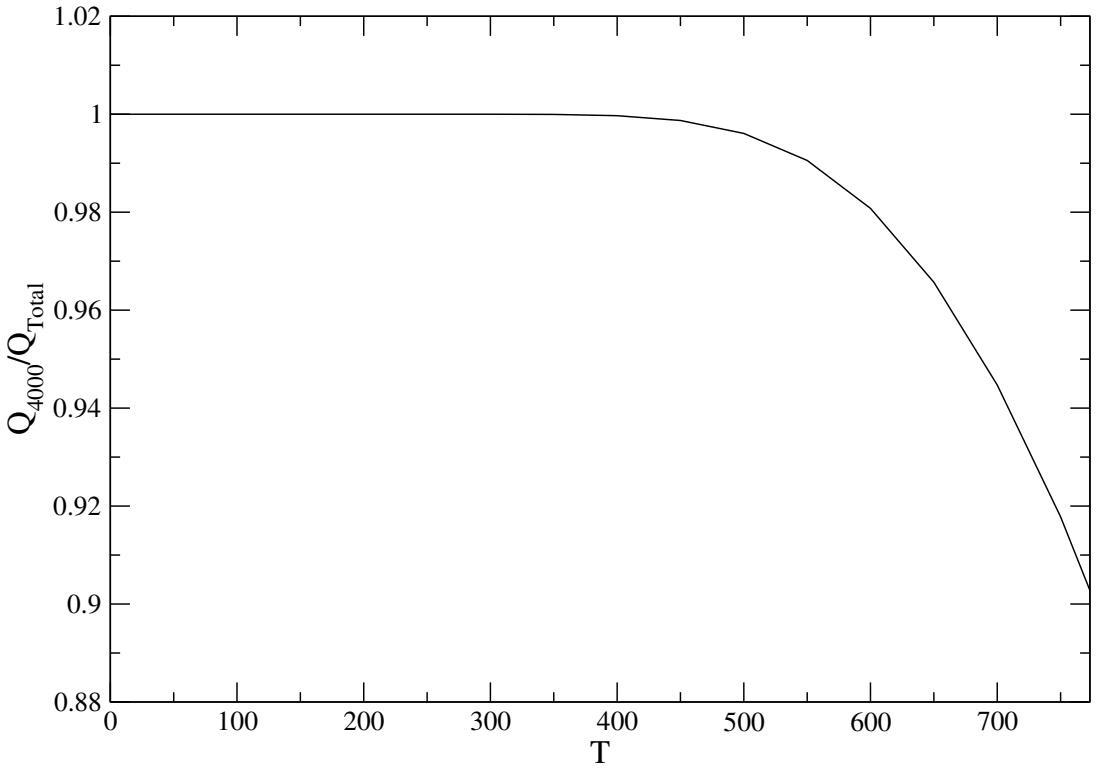

Figure 6.5: Ratios of Q_{4000} to the assumed converged values Q_{Total} as a function of temperature.

At T = 773.15 K, the line list is roughly 90% complete. In reality, this is an upper limit due to the fact that there is a slight under estimation in Q_{Total} at this temperature. However, the contribution from energies above 4000 cm^{-1} for any missing $J > 130$ levels can be considered negligible enough not to affect this beyond more than 1%.

6.5 Intensity Simulations

With a fully converged partition function at the given temperature a synthetic spectrum may be computed as a line intensity plot, via Eq. 4.11 of Chapter 4.

6.5.1 Comparisons with Relative Intensity Measurements

Chapter 4 provides a room-temperature (T = 298.15 K) comparison between UYT and the experimental data available. Here, the same comparisons are made using the UYT2 line list. Table 6.8 contains various details of the experimentally studied bands,

demonstrating the ability of UYT2 to reproduce the experimental band positions and the band intensities calculated in UYT.

Table 6.8: Comparison of calculated band intensities between UYT and UYT2 at T = 298.15 K for bands with experimentally recorded intensities (see Table 4.7 in Chapter 4). Units are given in 10^{-18} cm molecule^{-1}.

Band	Band Intensity	
	UYT	UYT2
$2\nu_2$ - ν_2	0.66	0.62
ν_2 - ν_0	3.71	3.39
ν_2 + ν_4 -ν_4	0.58	0.54
ν_4 - ν_0	5.95	5.37
$2\nu_4^{(l_4=0)}$ - ν_4	0.41	0.44
ν_2 + ν_4 - ν_2	0.53	0.49
$2\nu_4^{(l_4=2)}$ - ν_4	0.87	1.17
ν_1 - ν_4	0.10	0.22
ν_3 - ν_0	44.44	43.21
$2\nu_3^{(l_3=2)} - \nu_0$	0.12	0.11

Figures 6.6 - 6.8 show some detailed intensity comparisons for various bands at room temperature. The theoretical UYT2 data produced by TROVE is compared to a number of bands studied in the works of Maki *et al.* [99–105], in the same vein as Figure 4.5 in Chapter 4. The UYT2 data show an improvement in the band positions due to the PES refinement discussed earlier in this chapter. The absolute intensities of the experimental data are scaled via the same method as in Chapter 4.

There is no existing absolute intensity data for $^{32}S^{16}O_3$ above T = 298.15 K in the published literature, therefore making it difficult to assess the quality of the *ab initio* DMS used in the intensity simulations. Discussion is limited to that of Chapter 4 in which relative band intensities are compared between the observed and calculated data. Despite these available comparisons, there still remains ambiguity in both the relative intensities recorded across different pressures and the reproduction of intensities by the DMS.

6.5.2 Comparisons with Cross Section Data

Since the comparison with the data of Maki *et al.* is only able to provide a measure of the quality of relative intensities within a particular band, there is only one remaining method of assessing the overall DMS quality in the absence of available absolute inten-

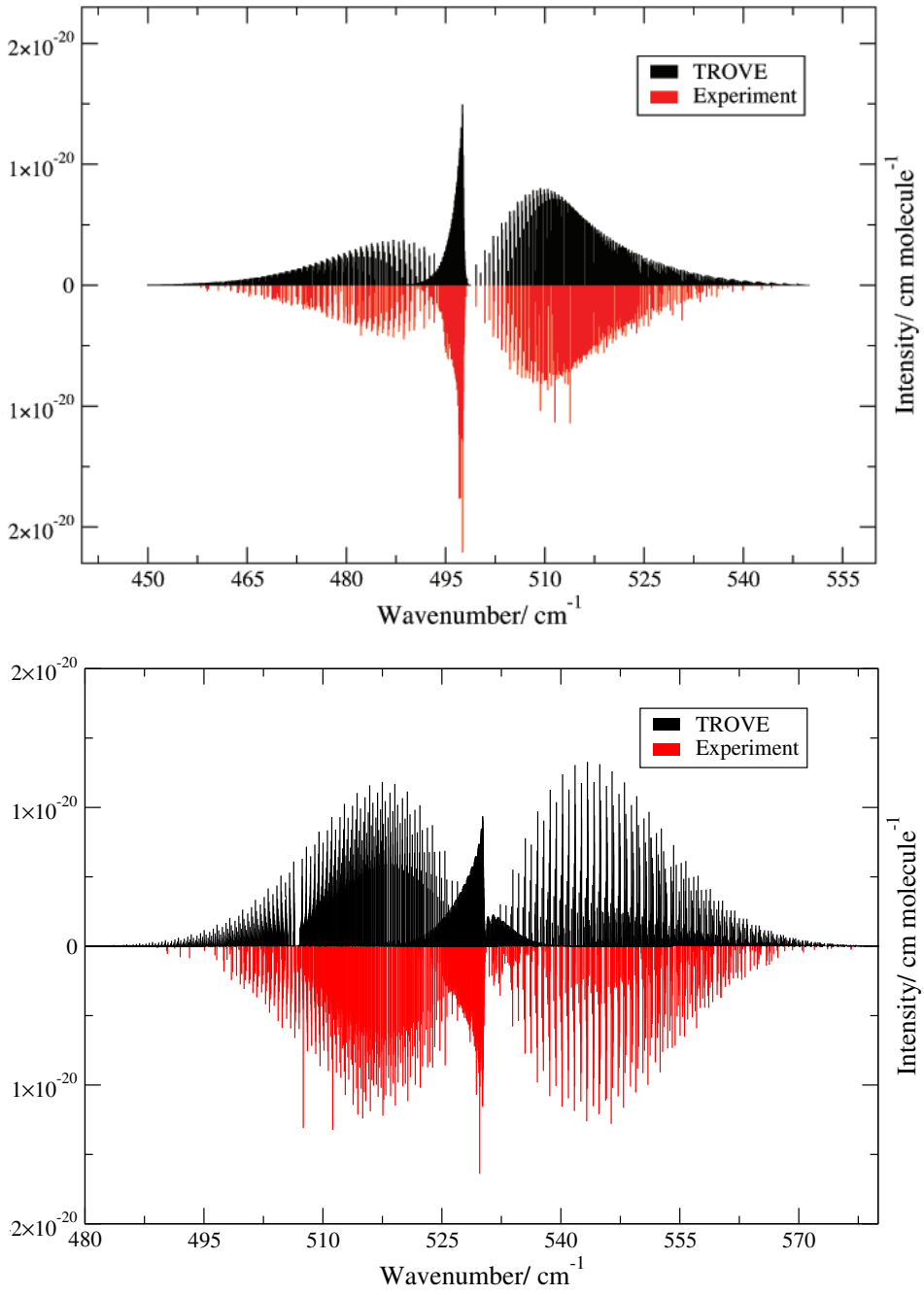

Figure 6.6: Comparison of the ν_2 - ν_0 (above) and ν_4 - ν_0 (below) bands at T = 298.15 K, between UYT2 data produced using TROVE and experimental data recorded by Maki *et al.* [99–105] at 0.409 Torr.

sity measurements, and that is through the generation of cross sections. During the calculation of the hot line list a handful of experimentally derived SO_3 cross section data were made available for a range of temperatures by Alexander Fateev at DTU [187], including data recorded by PNNL at room-temperature [126]; for both data sets there are discernible features across four separate spectral regions. As with SO_2 in Chapter

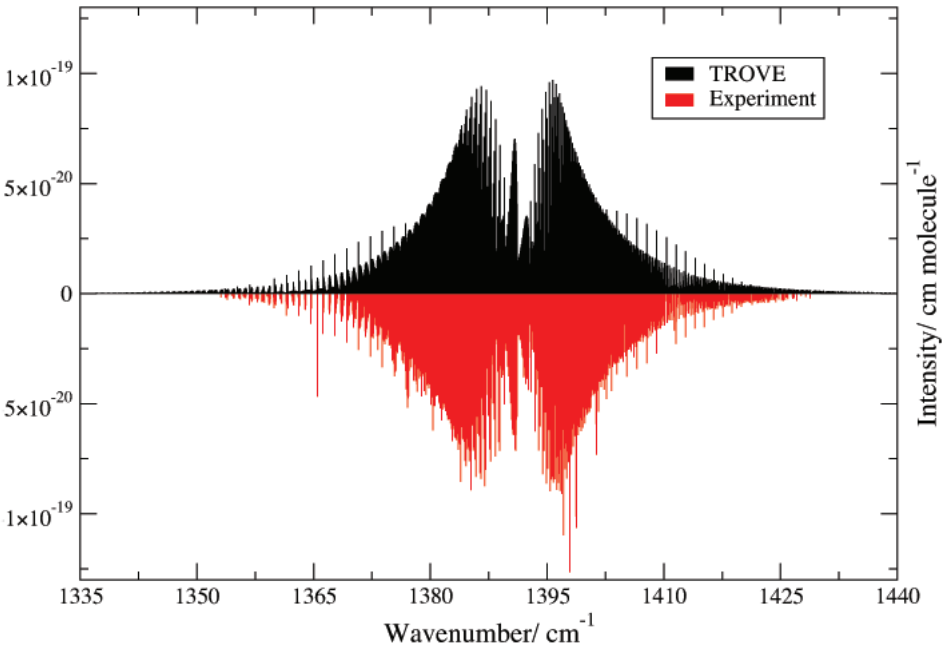

Figure 6.7: Comparison of the ν_3 - ν_0 band at T = 298.15 K, between UYT2 data produced using TROVE and experimental data recorded by Maki *et al.* [99–105] at 0.16 Torr.

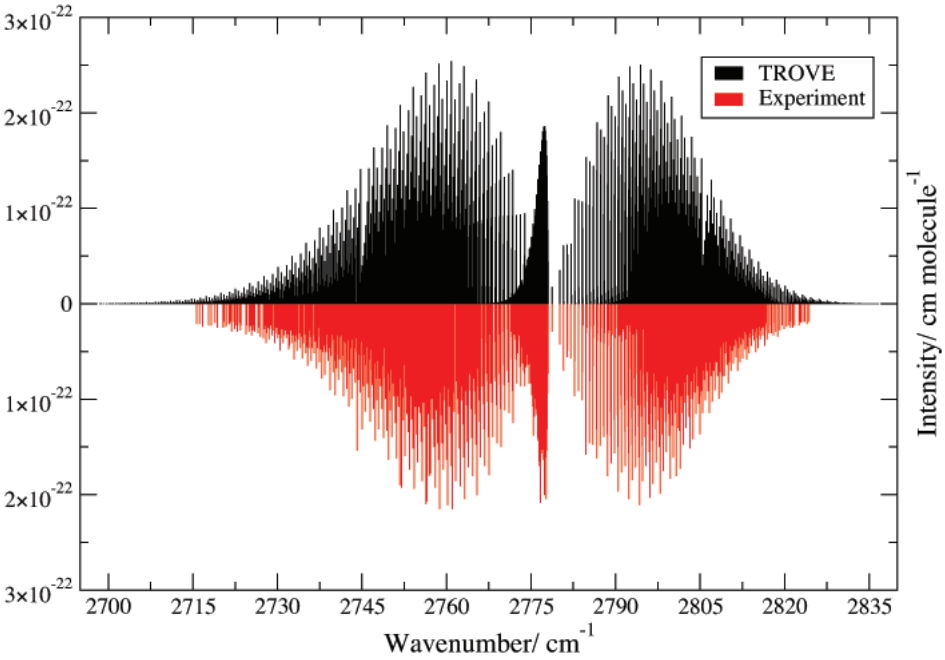

Figure 6.8: Comparison of the $2\nu_3$ - ν_0 band at T = 298.15 K, between UYT2 data produced using TROVE and experimental data recorded by Maki *et al.* [99–105] at 4.99 Torr.

3 it should be possible to make a semiquantitative analysis by comparing integrated intensities across a given spectral window. To this end, the ExoCross tool by Yurchenko is used to produce a number of theoretical, temperature-dependent cross sections, as in Chapter 3 for SO_2.

The difficulty of obtaining the full UYT2 line list made simulating accurate cross sections practically impossible until all energies and dipole transitions were computed. It would be possible to use UYT to make such comparisons at room-temperature, however line-position inaccuracies and possible ambiguity of the computed partition function would make it hard to draw a valid conclusion. Figures 6.9 and 6.10 show comparisons between recorded cross sections from PNNL at 298.15 K (25°C) compared with simulated cross sections using the full UYT2 line list, based on a Gaussian profile of HWHM = 0.1 cm^{-1}.

Figure 6.11 shows a comparison of the the ν_2 and ν_4 complex and the ν_3 band between cross sections recorded for SO_3 at 573.15 K (300°C) and those simulated using the UYT2 line list, based on a Gaussian profile of HWHM = 0.25 cm^{-1} (this difference in HWHM value across different temperatures is mainly based on the best representation of the experimental spectra; in actuality the integrated intensity across the spectral region is largely independent of the HWHM value used). Figure 6.12 shows the same for the $\nu_1 + \nu_3$ band, which appears in a noisier region of the spectrum, and is also disturbed by a strong, foreign absorption feature resulting from the presence of CO_2. There is no data at T = 573.15 K for the $2\nu_3$ band due to noise contamination in the associated spectral region. Measurements of SO_3 have also been made for 773.15 K (500°C), however it has not been possible to generate accurate experimental cross section values due to difficulties in estimating the concentration within the gas flow cell; it is believed that at this temperature a significant amount of the SO_3 gas is prone to reacting with material within the measurement apparatus [187]. Therefore, this data is not included in the current comparison.

As can be seen there is a clear tendency for the UYT2 data to overestimate the line intensities for both temperatures, though band positions and features are qualitatively fairly well represented. There are a number of possibilities that could explain these discrepancies. Firstly, it is possible that the experimental cross sections may be underestimated due to an error in the SO_3 abundances; the calculation of cross sections requires the knowledge of the species concentration within the length of the absorption cell [244]. The fact that measurements at room-temperature by Fateev *et al.* corroborate the PNNL data, and that similar discrepancies are observed for both data sets suggests that this is not the case [187].

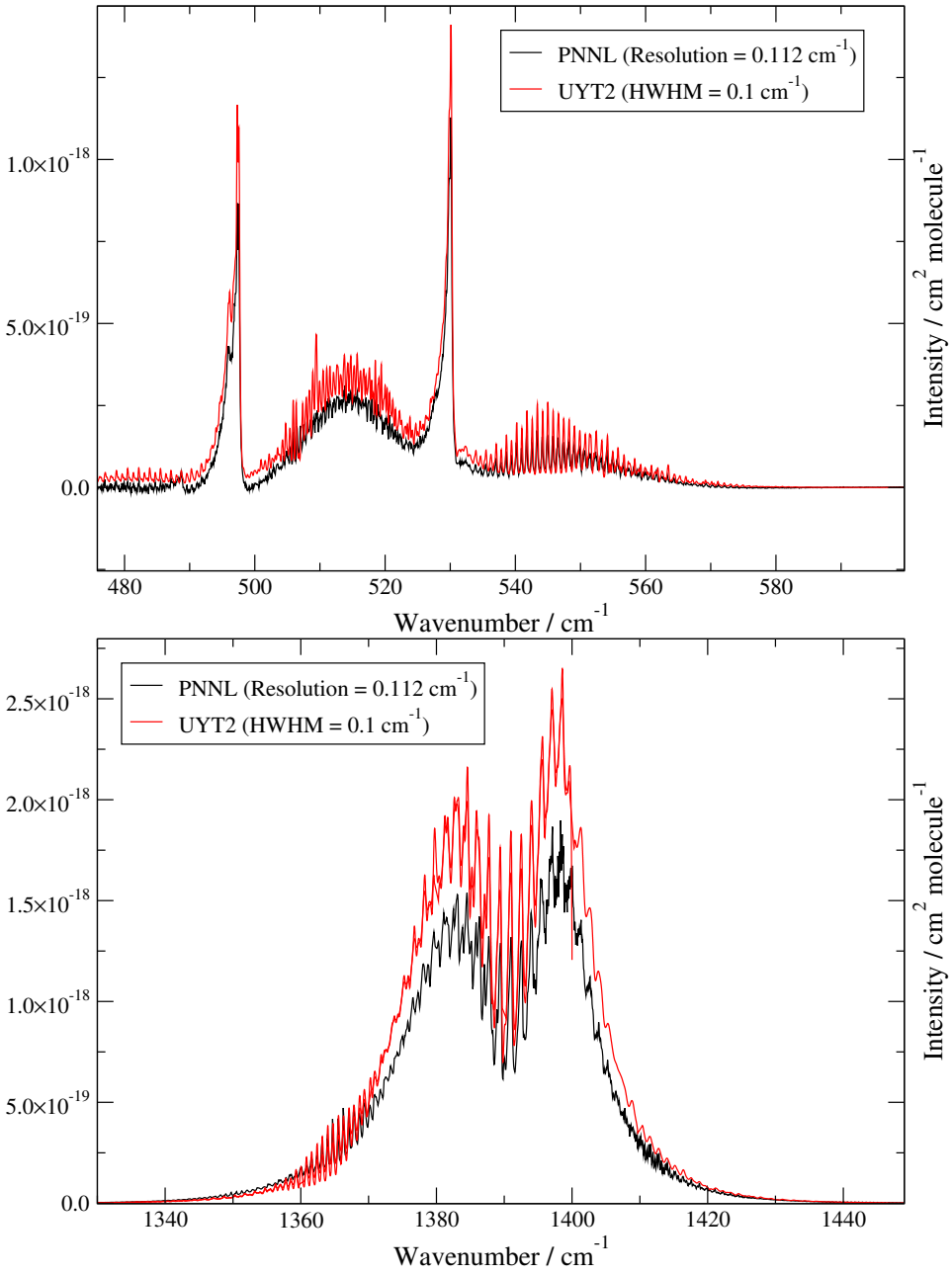

Figure 6.9: Comparisons of the ν_2 and ν_4 bands (above) and the ν_3 band (below) for experimentally obtained [126] and simulated cross sections at T = 298.15 K.

The second possible source of disagreement can be attributed to convergence issues in the partition function. Since the calculated intensities given by Equation 2.55 depend on the scaling factor $Q(T)$ the incorrect computation of this value at the given temperature will lead to inaccurate values of absolute intensity. The difference in integrated cross section intensities observed suggest that, if the calculated value of $Q(T)$ is incorrect, then it is smaller than the 'true' value, since the theoretical cross sections are more intense than the experimentally observed values. This scenario can also be ruled

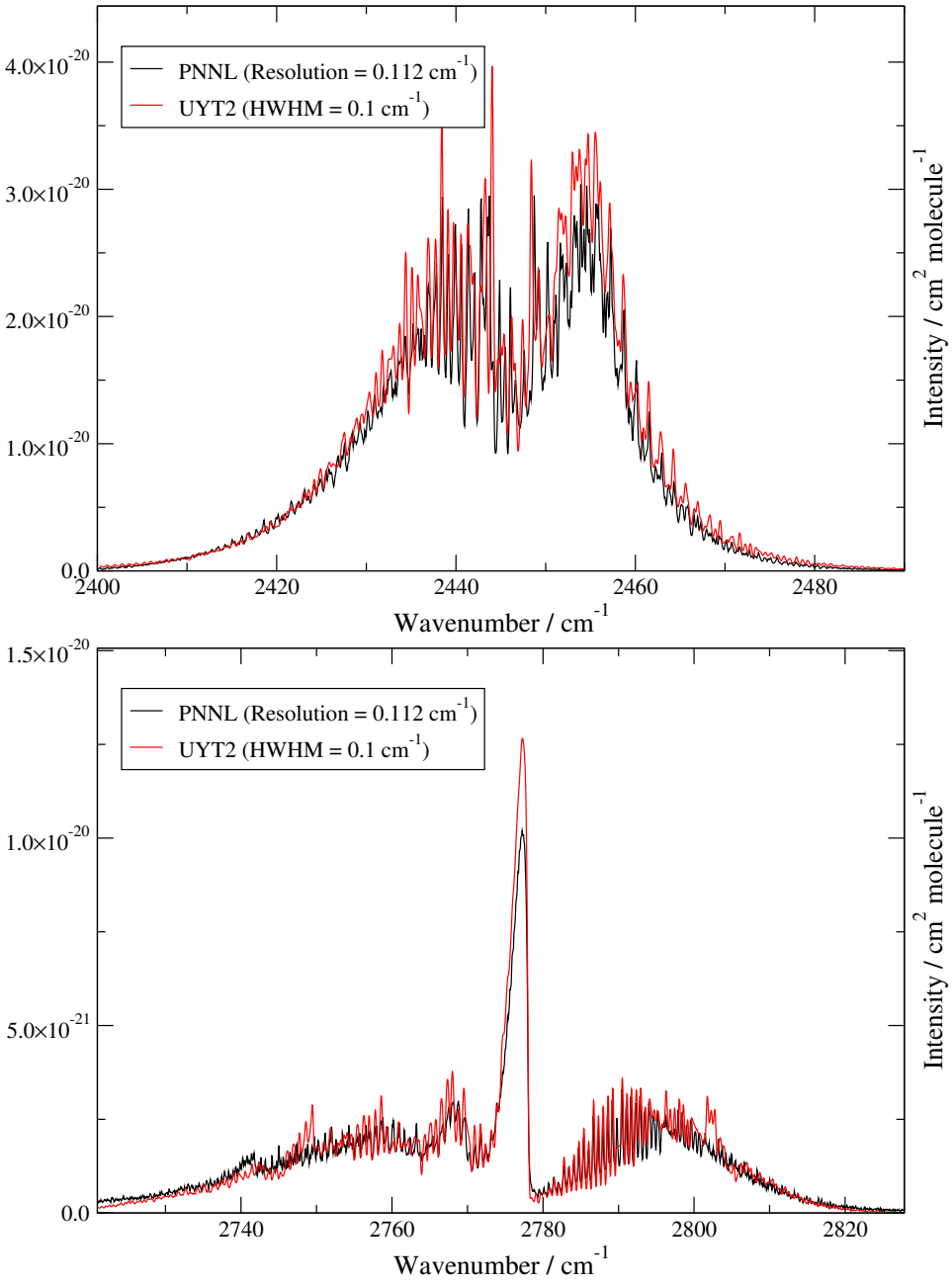

Figure 6.10: Comparisons of the $\nu_1 + \nu_3$ bands (above) and the $2\nu_3$ band (below) for experimentally obtained [126] and simulated cross sections at $T = 298.15$ K.

out, due to two reasons. First, the agreement between $Q(T)$ for both UYT and UYT2 is very good at $T = 298.15$ K, where they are both adequately converged; the increased basis set size underlaying the UYT2 calculations would undoubtedly account for any missing rovibrational energies in UYT. Secondly, and perhaps more interestingly, the analysis of several bands across different temperatures shows the cross section discrepancies to be almost independent of the value of T (see below). This would not be expected if $Q(T)$ were the source of the disagreement, since it has a nonlinear depen-

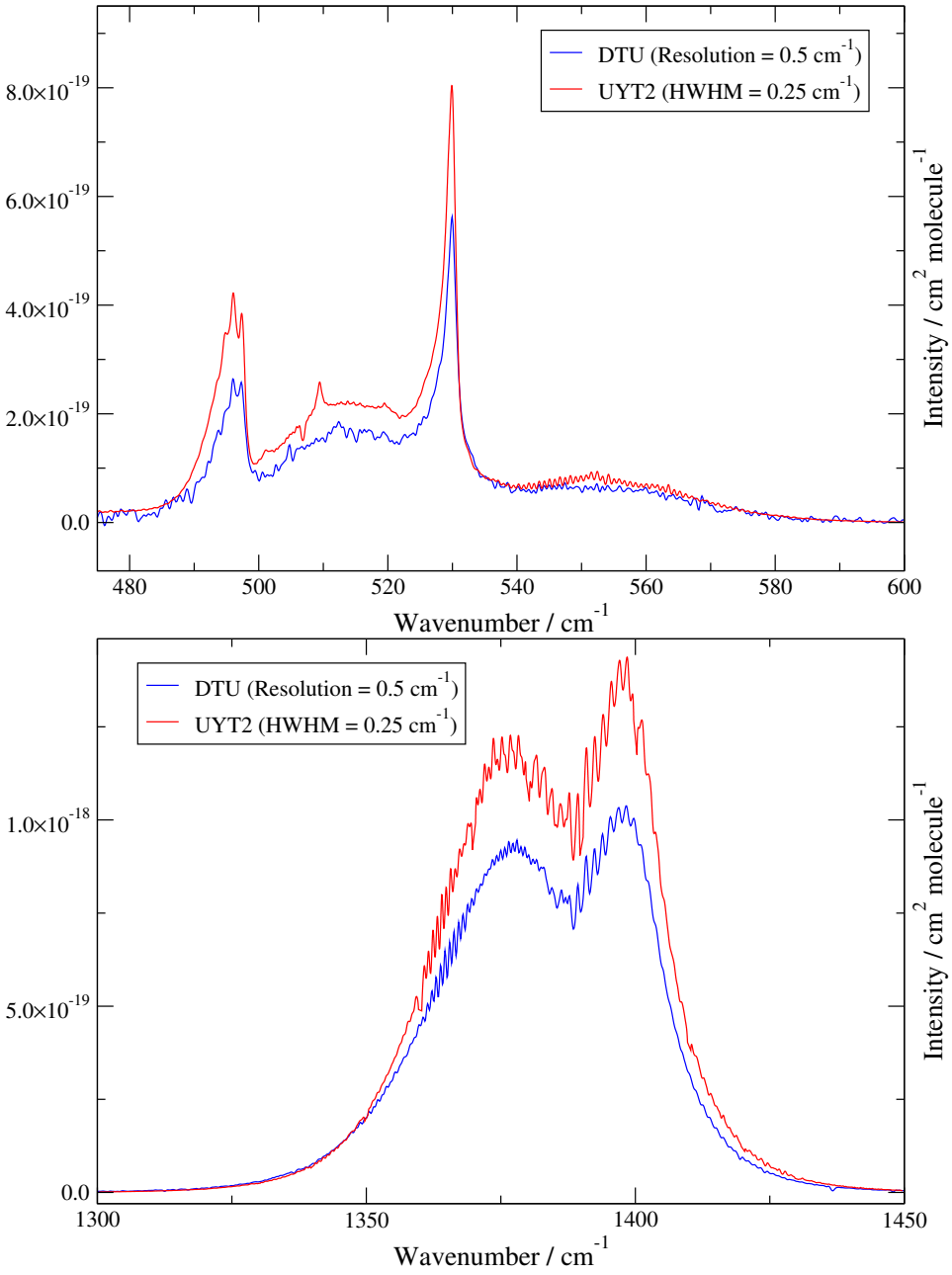

Figure 6.11: Comparisons of the ν_2 and ν_4 bands (above) and the ν_3 band (below) for experimentally obtained [187]and simulated cross sections at T = 573.15 K.

dence on the temperature.

This heavily implies that the problem lies with the DMS. The successful reproduction of SO_2 cross sections using the ExoCross tool by Yurchenko can rule out any computational defects in this procedure, and therefore the error may be attributed to the final computed values of Einstein A coefficients for SO_3. DMS defects have been observed previously within the ExoMol group [171, 245, 246], despite previous experience of obtaining accurate *ab initio* dipole surfaces; in general, it is highly beneficial

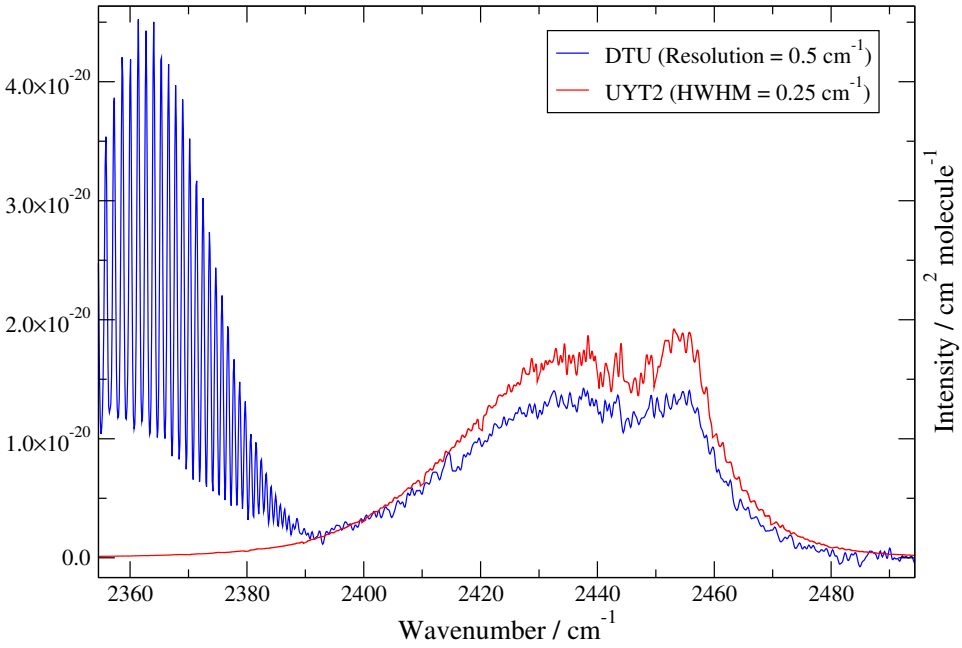

Figure 6.12: Comparisons of the $\nu_1 + \nu_3$ bands (above) and the $2\nu_3$ band (below) for experimentally obtained [187] and simulated cross sections at $T = 573.15$ K.

to obtain accurate absolute intensity measurements in order to validate any computed DMS, which is unfortunately not the case for SO_3.

It is not easy to make a rigorous analysis based on cross section data available for SO_3, as it is not immediately obvious what the contributions are from individual lines. In addition to this, both data sets contain varying degrees of noise within certain spectral regions, with the region around the ν_3 band generally providing the best signal. Pavlyuchko *et al.* in ref. [246] performed a fit of their DMS based on experimental intensity data for nitric acid, to better improve simulated intensities. The lack of absolute intensity measurements for SO_3, coupled with the expensive computational demands of the line list calculation make this particularly difficult to perform here. Nevertheless, the best approach has been to compare integrated band intensities across fixed spectral windows to obtain scaling parameters for the each band. Table 6.9 summarises the ratios of integrated intensities between simulated and recorded cross sections for some available bands.

For most bands there appears to be a fairly consistent shift in intensity values across different temperatures, however the overtone bands for $T = 298.15$ K suggest otherwise. The differences are quite subtle; for example, while the $2\nu_3$ band has almost perfect agreement in integrated intensity across the band, the central Q-branch peak is

Table 6.9: Integrated band intensities calculated from observed and calculated (UYT2) cross sections. Intensity units are given in 10^{-18} cm^2 molecule^{-1}.

Temperature (K)	Band	Integrated Band Intensity		Obs./UYT2
		Obs.	UYT2	
298.15	ν_2 & ν_4	9.95	13.13	0.76
	ν_3	46.78	60.38	0.77
	$\nu_1 + \nu_3$	0.71	0.82	0.87
	$2\nu_3$	0.15	0.16	0.97
573.15	ν_2 & ν_4	10.26	13.53	0.76
	ν_3	46.79	59.62	0.78
	$\nu_1 + \nu_3$	0.69	0.87	0.79

not well represented by the UYT2 cross sections. On the other hand, the data at 573.15 K by Fateev *et al.* for the $\nu_1 + \nu_3$ band exhibits the same general shift as the ν_2, ν_3 and ν_4 bands (care has been taken to not include the intense contamination in the integration), while the same is not true at room-temperature. The PNNL room-temperature cross sections are presented as a composite spectrum created from 8 individual absorbance spectra taken at various different pressures using both a mid-band MCT and wide-band-MCT detector [126], and uncertainties in intensity are listed as 10%. Conversely, the high-temperature data is a continuous spectrum which, while being prone to noisier regions, still provides some measurement consistency. If the scaling factors for the two overtone bands at room-temperature are ignored, then the remaining factors may be averaged (0.76) and applied to all simulated cross sections. The assumption made here is that the apparent stronger agreement in the room-temperature intensities for the $\nu_1 + \nu_3$ and $2\nu_3$ are 'accidental', while the wide, coverage-consistent high-temperature cross sections provide a more accurate description of the differences. It should be noted that defective DMSs are usually more likely to present non-consistent differences in intensity across different bands [171, 245, 246]. Without extra experimental data for more bands at different temperatures it is difficult to ascertain whether the same is true in this case.

Figures 6.13 and 6.14 shows the various bands at room-temperature, with computed cross sections multiplied by the averaged scaling factor. Figures 6.15 and 6.16 show the same for T = 573.15 K, which improves the simulated cross sections, and demonstrates the implied temperature independent nature of the discrepancy. As can be seen in Figure 6.14, using the averaged scaling factor (obtained from excluding the

individual $\nu_1 + \nu_3$ and $2\nu_3$ Obs./UYT2 ratios) improves the reproduction of the central band peak, though the P-branch does show some intensity differences. This appears to be common for multiple bands and is possibly due to the neglecting pressure broadening (see Chapter 7).

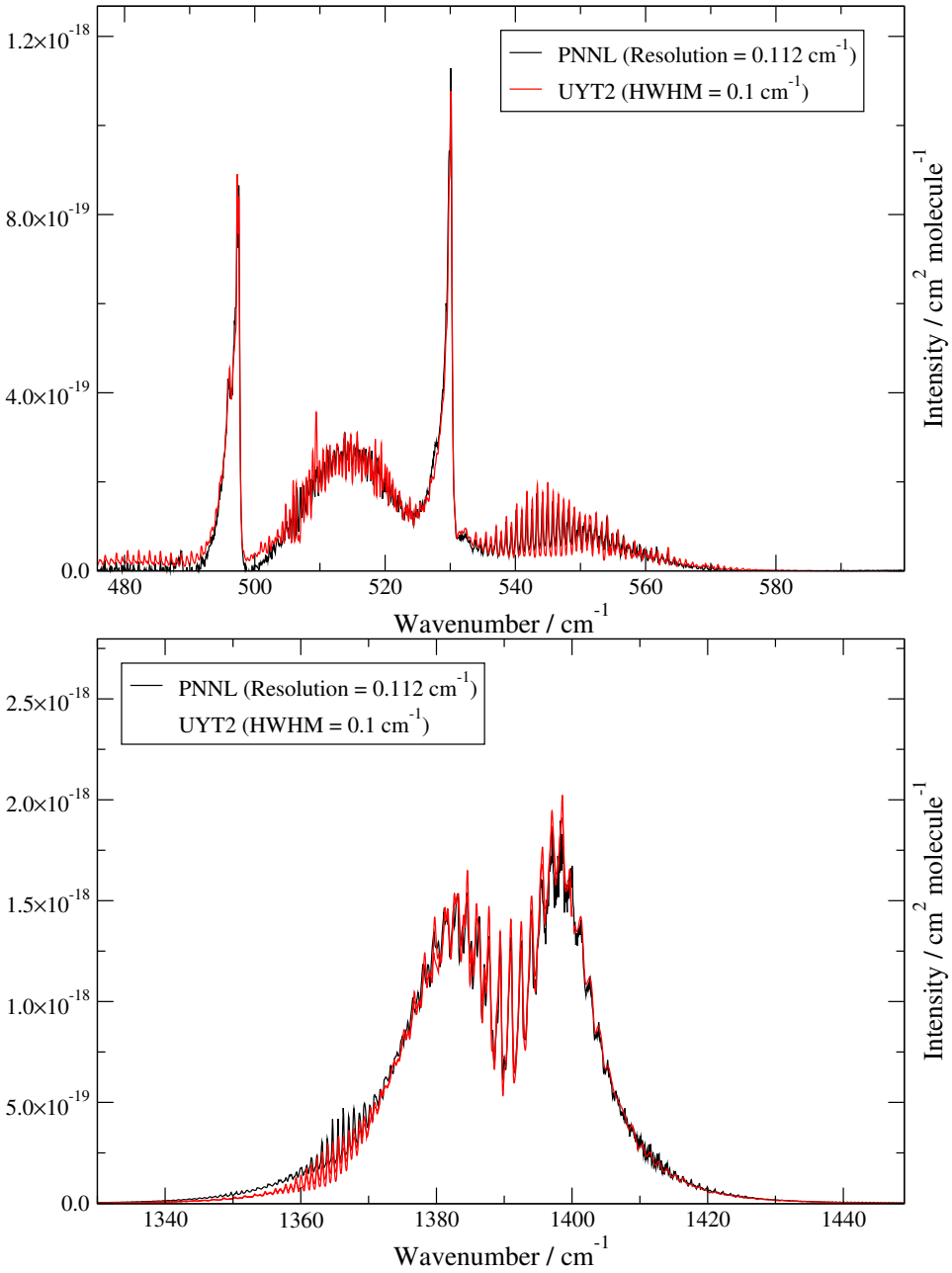

Figure 6.13: Comparisons of the ν_2 and ν_4 bands (above) and the ν_3 band (below) for experimentally obtained [126] and scaled simulated cross sections at T = 298.15 K.

Figure 6.17 shows the cross sections calculated over the entire spectral range of $0 < \nu \leq 5000$ cm^{-1}, using a Gaussian profile of HWHM = 0.25 cm^{-1}, for a number of different temperatures. All simulated cross sections have been multiplied by the

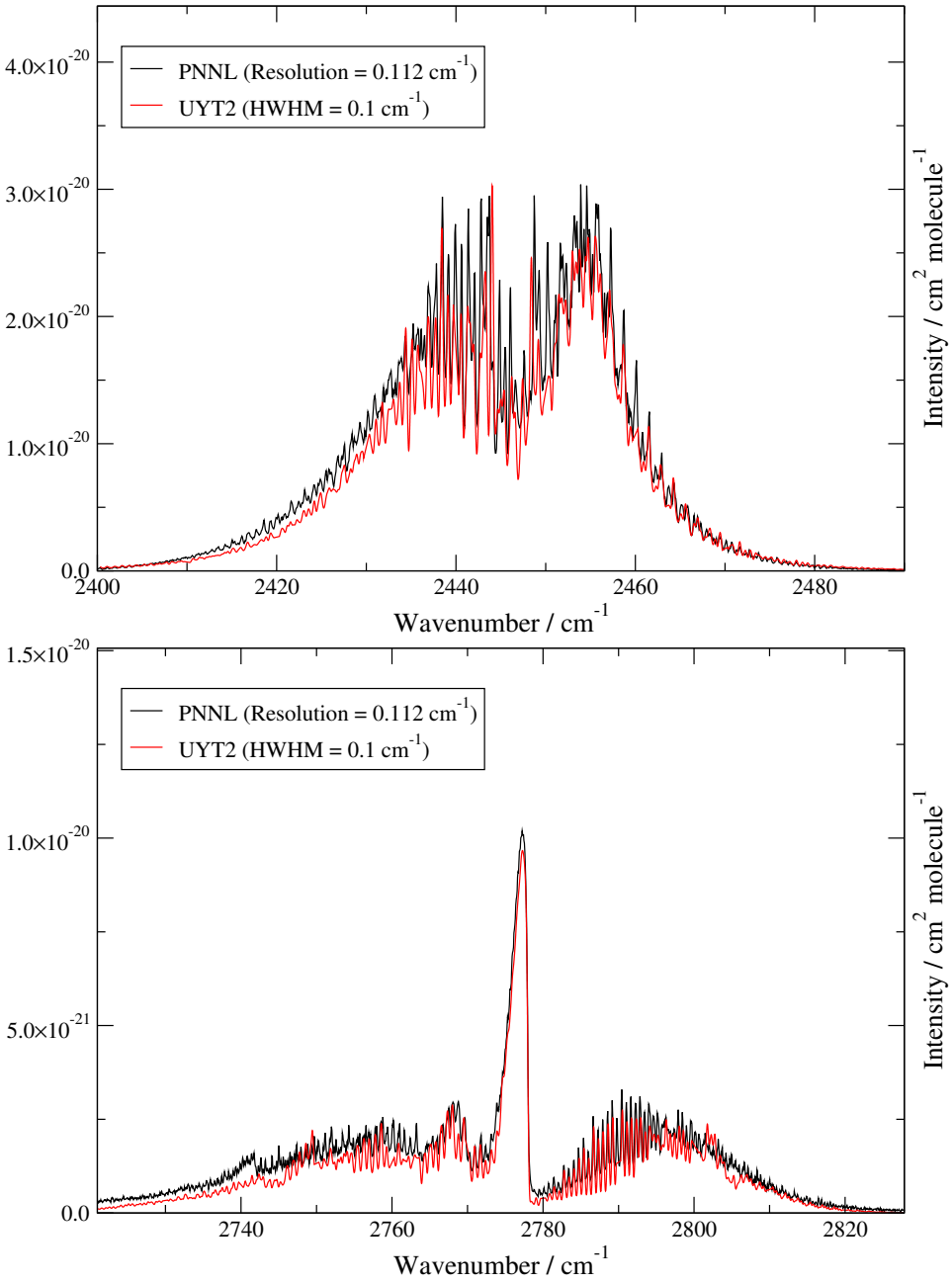

Figure 6.14: Comparisons of the $\nu_1 + \nu_3$ bands (above) and the $2\nu_3$ band (below) for experimentally obtained [126] and scaled simulated cross sections at $T = 298.15$ K.

average scaling factor discussed.

6.6 Discussion

The UYT2 line list contains 21 billion transitions, and a total of 18 million energy levels below 10 000cm^{-1}. This provides an improvement upon the initial room-temperature line list, UYT, in terms of both line positions and temperature coverage.

Table 6.2 provides a measure of the improvement introduced by the PES refine-

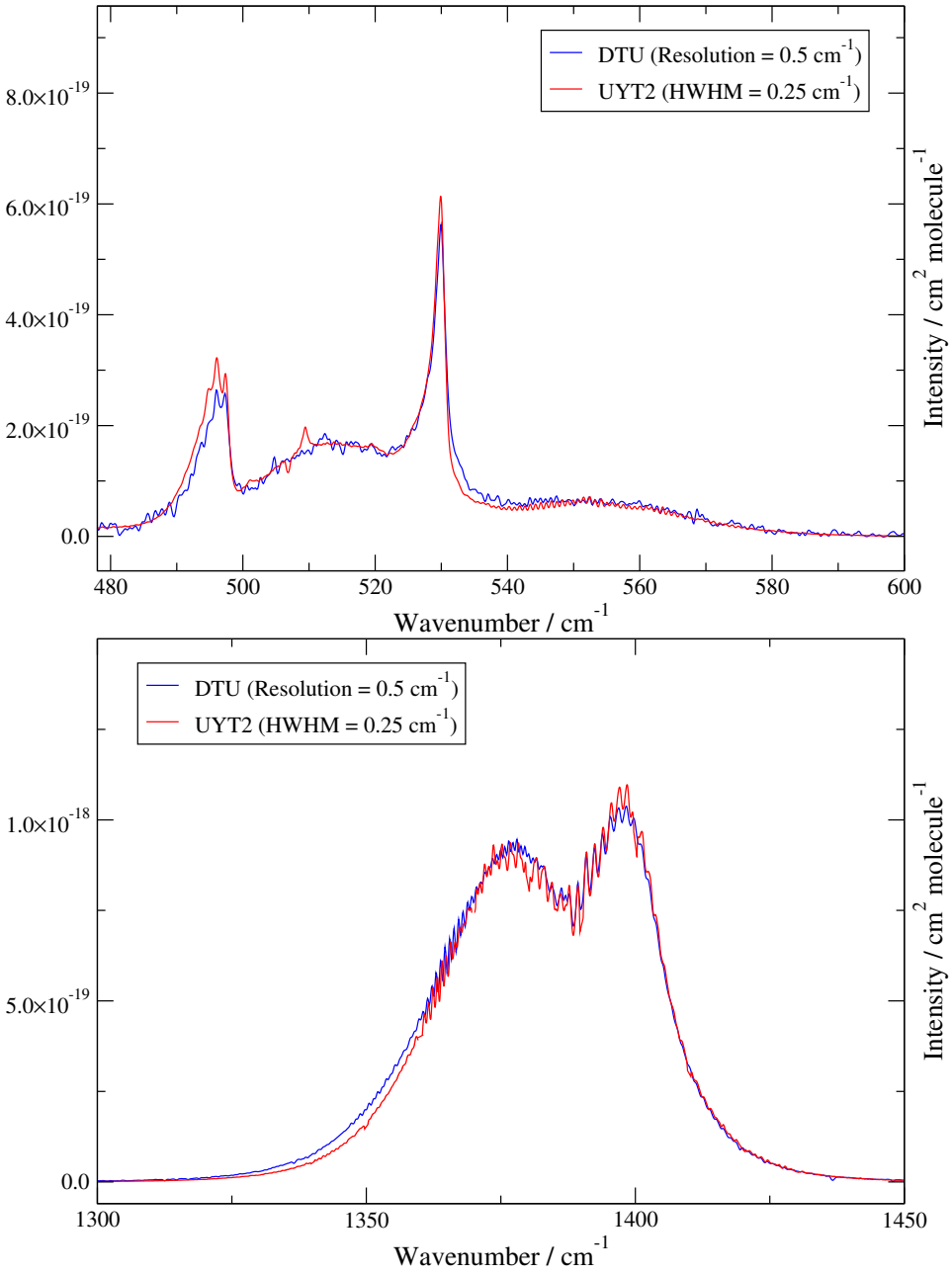

Figure 6.15: Comparisons of the ν_2 and ν_4 bands (above) and the ν_3 band (below) for experimentally obtained [187] and scaled simulated cross sections at T = 573.15 K.

ment present in the UYT2 line list. The total RMS deviation for the bands included in the potential adjustment is 1.35 cm^{-1}, compared to 3.23 cm^{-1} for the unrefined PES of UYT. The majority of simulated line positions across these bands is improved by an order of magnitude.

As with the UYT line list described in Chapter 4, it is difficult to ascertain the overall quality of the *ab initio* DMS used in the production of line intensities. However comparing with newly available cross section data at two different temperatures heavily

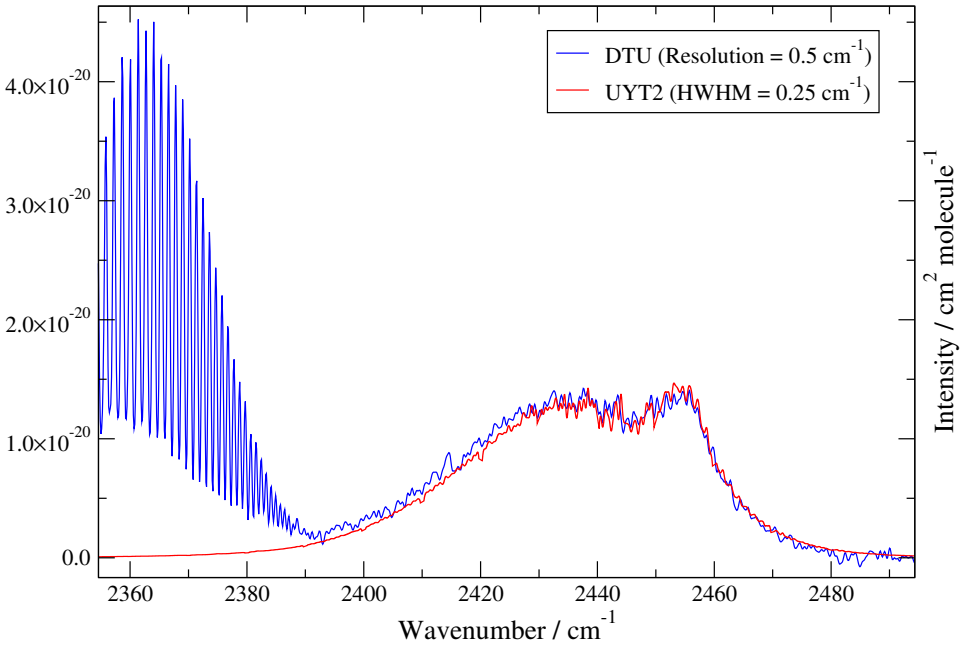

Figure 6.16: Comparisons of the $\nu_1 + \nu_3$ band for experimentally obtained [187] and simulated cross sections at $T = 573.15$ K.

suggests that the DMS used in the calculation of intensities is slightly defective, causing an apparently constant shift in all intensity values. The evidence suggesting this temperature- and band-independent scaling factor is certainly not conclusive, and one may wish to take care in which scaling factor to use for each band. In particular, bands for which there are no experimental intensity data available can not be considered to be truly represented well in UYT2 and the lack of exhaustive absolute intensity knowledge for SO_3 at the current time of writing severely limits the ability to effectively correct for the disagreements observed.

For the purposes of discussion, it is interesting to observe what effect that different levels of theory have on resulting *ab initio* dipole moment calculations; it is possible that the DMS will be highly dependent on the use of basis set in the electronic structure calculations. Table 6.10 shows the results of CCSD(T) [115] calculations using varying basis sets for a few nuclear geometries, using Molpro [192]. Included are calculations using a standard triple-zeta basis set [150], those optimised for F12 calculations [191], as used in this work (and used in conjunction with the appropriate method [190]), and finally with the quadruple-zeta basis set [184, 185] used in the construction of the SO_2 DMS of the current work [35] (see Chapter 3). There is no experimentally measured dipole moment data for SO_3 as it does not posses a permanent dipole moment, and

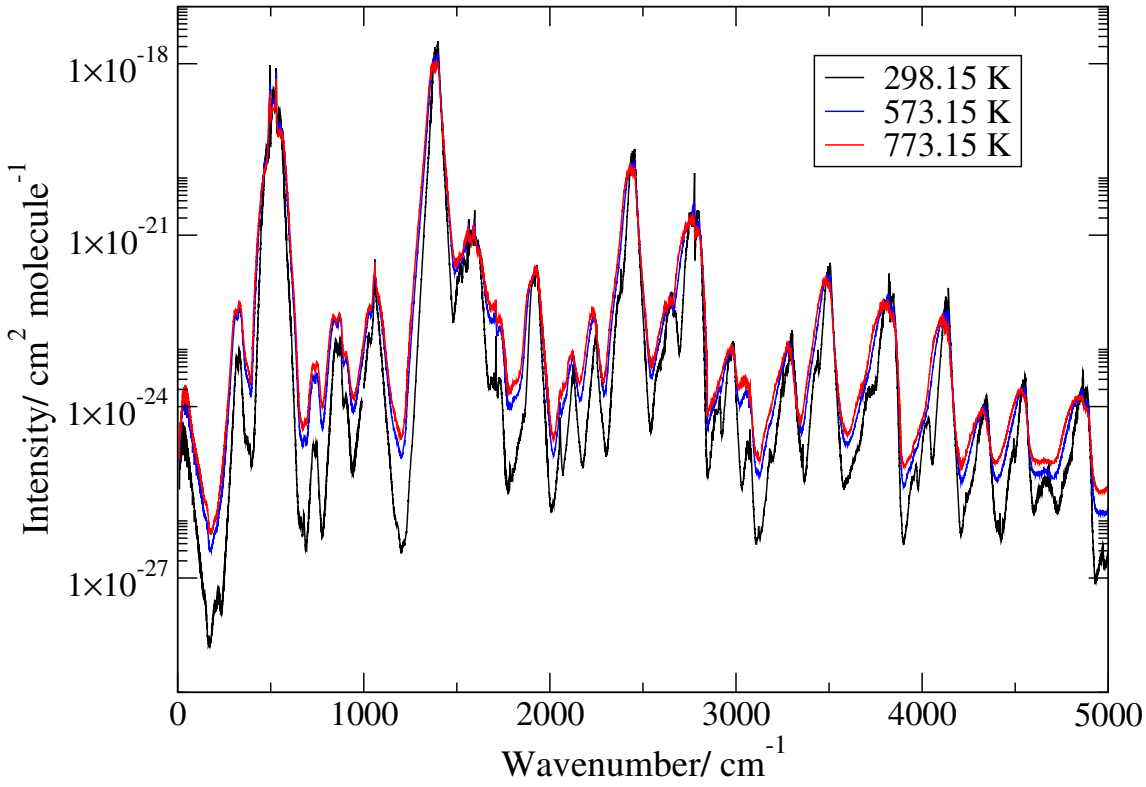

Figure 6.17: Overview of the simulated cross sections using UYT2, at T = 298.15, 473.15 and 773.15 K, with a Gaussian profile of HWHM = 0.25 cm^{-1}.

therefore points are chosen arbitrarily, one to reflect a symmetric geometry, and another for an asymmetric geometry. For these calculations the planar molecule lies within the yz-plane, and the component of the dipole moment is measured across both these axes for the instantaneous geometry.

Table 6.10: Components of dipole moment for a select number of nuclear geometries, computed a various levels of theory. Both configurations align the S-O_1 bond parallel to the z-axis. The symmetric configuration consists of an elongated S-O_1 bond of 1.517 Å, with the remaining bond lengths and angles at equilibrium (1.417 Å, and 120°, respectively; only the z-component of the dipole moment is present. The asymmetric configuration also has an S-O_1 bond length of 1.517 Å, with the S-O_2 and S-O_3 bonds at 1.417 Å and 1.317 Å, respectively, thus there is a dipole component across both the y and z-axes.

Level of Theory	Symmetric Geometry	Asymmetric Geometry	
	μ_z / D	μ_z / D	μ_y / D
CCSD(T)/aug-cc-pVTZ	0.3772	0.6004	0.3664
CCSD(T)-F12b/aug-cc-pVTZ-F12b	0.3785	0.6022	0.3669
CCSD(T)/aug-cc-pV(Q+d)Z	0.3758	0.5989	0.3652

It can be seen that the value of the dipole moment is somewhat sensitive to the choice of theory and basis set used, though in the absence of a fully-constructed DMS

it is difficult to say what effect this would have on the quality of *ab initio* transition intensity calculations. The DMS constructed in this work is based on an analytical fit to a grid of 13 000 geometry points, which reproduces the *ab initio* data with an rms of 0.00013 D (see Chapter 4). Therefore, though the deviations present in 6.10 are relatively small, they are larger than this quoted rms, which suggests that an equally accurate fitting to a set of higher-level *ab initio* points may produce a DMS which is structurally different to the one used in the current work, perhaps enough to have a significant effect on computed transition intensities. Investigating this would require a further systematic study into the dipole moment, like that carried out by Azzam *et al.* for H_2S [245], in which *ab initio* dipole calculations are extended to a large number of geometries in order to produce several DMSs based on different underlying basis sets, in order to test each of their effects on the calculation of transition moments. However the higher-dimensionality of the SO_3 problem, coupled with the lack of experimental dipole data for this molecule makes this a daunting task. Nevertheless it is hoped that the scaling factor of the cross-sections can improve the *ab initio* intensity values produced in the UYT2 line list, although further work will be required in order to fully investigate the source of the discrepancy. In the meantime the scaling factor can provide a lower limit to the estimated uncertainty in simulated cross section values, which is 31%.

The increased size of the basis set, the computation of rovibrational energies up to $J = 130$, and the increased spectral range of line strength calculations allows for UYT2 to be used in the simulation of spectra between $0 < \nu \leq 5000$ cm^{-1}, with approximately 90% completion at $T = 773.15$ K (500 C). Given that this is the largest data set of its kind for $^{32}S^{16}O_3$, it is recommended that UYT2 be used in the production of cross sections at room-temperature, and up to $T = 773.15$ K, for both astronomical and industrial applications. The full line list can be downloaded from the ExoMol website (http://www.exomol.com).

Chapter 7

General Discussions & Conclusions

The main focus of this work has been in the production of highly accurate infrared line lists for the main isotopologues of SO_2 and SO_3. These have been computed via the use of spectroscopically-refined *ab initio* ground-state potential energy surfaces (PES) and *ab initio* dipole moment surfaces (DMS), in conjunction with bespoke nuclear motion codes.

The SO_2 line list is based upon a CCSD(T) [115] calculated PES using a cc-pVQZ-DK basis set [184], refined by the use of 30 129 experimentally obtained energy levels [35]. The DMS is computed at the same level of theory with an aug-cc-pV(Q+d)Z [184, 185] basis set. Both the PES and DMS have been implemented into the DVR3D program suite [138] which has been optimised for the purposes of computing high-J rovibrational energy levels and dipole transition strengths. The final line list contains 1.3 billion transitions within the $0 < \nu < 8000$ cm^{-1}spectral region between 3 255 954 rovibrational energy levels up to $J = 165$ and $E' \leq 15\,000$ cm^{-1}. Analysis of the partition function suggests that this line list should be suitable for simulating rovibrational SO_2 spectra for $T \leq 1500$ K, and may be used approximately for higher temperatures. The agreement with the synthetic line list produced by Huang *et al.* [35] is good, with a standard deviation of 0.043 cm^{-1} in matched lines, with a corresponding agreement to within 2% for absolute intensities. Agreement with HITRAN [75] is also very good, with an overall root-mean-square error of 0.016 cm^{-1} in line positions, and within 80% - 90% for absolute intensities. Preliminary comparisons with cross section data at 573.15 K (300 °C) and 773.15 K (500 °C) shows good semiquantitative agreement, with 2% and 6% difference, respectively, in integrated intensities across the ν_1 and ν_3 fundamental bands.

The SO_3 line list is based upon a CCSD(T)-F12b [190] calculation for the PES and DMS, using aug-cc-pVTZ-F12 and aug-cc-pV(T+d)Z-F12 [191] basis sets for oxygen and sulphur, respectively. The *ab initio* PES and DMS were used to create an initial room-temperature line list using the TROVE program [131], consisting of 174 674 257 transitions within the 0 - 4000 cm^{-1} region between 4 814 799 rovibrational energy levels up to $J = 85$ and $E' \leq 7000\ cm^{-1}$. This line list was compared to the experimental data in the literature [99–105] where it was deemed necessary for a refinement of the *ab initio* PES in order to correctly replicate band positions. With a lack of experimentally obtained absolute intensities it was only possible to semi-quantitively determine the accuracy of the DMS, which was able to replicate relative intensities fairly well for a number of bands.

This was followed by a refinement of the PES based on 119 reliable experimental energy levels, reducing the root-mean-square deviation from 3.23 cm^{-1} to 1.35 cm^{-1} over all compared bands, improving the majority of most band positions by an order of magnitude. This refined PES was used to compute a hot line list for SO_3, which contains 21 billion transitions between 18 million rovibrational energy levels below for $J \leq 130$ and $E' \leq 10\ 000 cm^{-1}$. Comparisons with experimental cross section data at 298.15 K and 573.15 K from separate sources have cast doubt on the accuracy of the *ab initio* DMS, where a temperature-independent difference in intensities has been observed between experimental and calculated data. Nevertheless, some preliminary analysis has been performed in an effort to scale the computed data to better replicate the observed intensities. Convergence of the partition function suggest that this line list is suitable for simulating spectra up to 773.15 K, but with possibility of approximating for temperatures above this.

Finally, the rotational dynamics of the SO_3 molecule have also been investigated from a theoretical perspective, where the source of the ‘forbidden’ rotational spectrum has been quantum mechanically analysed, and attributed to so-called ‘intensity stealing’ based on the centrifugal distortion of the molecule, attributed to the bending modes, which induces a dipole moment. In addition, energy cluster formation is predicted for this molecule for highly excited rotational states, providing a unique, first-time observation of such behaviour in a planer, tetratomic molecule containing no hydrogen.

7.1 Further Work

There are a number of implications for further work based on the studies presented in this thesis. There are a variety of applications which may benefit from usage of both the SO_2 and SO_3 line lists, across a multitude of scientific disciplines. There are also a number of areas of improvement which must be considered with regards to the results obtained from the current work.

7.1.1 The Hot SO_2 Line List

Since the hot SO_2 line list computed in the current work is based on the PES and DMS by Huang *et al.* the conclusions drawn in Refs. [35] and [118] and subsequent implications for future work are particularly relevant here. In particular, the need for accurate PES refinement for the purposes of simulations at higher energies and temperatures is discussed. It is hoped that the current SO_2 line list may facilitate in the analysis of high resolution IR spectra, particularly in spectral regions where experimental coverage is relatively low, and that assignments of lines within existing experimental datasets may in turn be used to further refine the PES to more accurately reflect the true behaviour in the higher energy regions. As has been previously stated, the rovibrational energies calculated using DVR3D are only exact within the limitations of the underlying PES, as they intricately dependent on its accuracy. The extension of wavenumber coverage in the rovibrational spectrum would require the determination of an accurate *ab initio* PES at higher energies. In a similar vein the simulation of high-temperature spectra will also enable validation of the DMS beyond 296 K, however early indications for two fundamental bands show good signs of semiquantitative agreement. It is hoped that further collaboration with Huang and coworkers will enable a better understanding of this molecule at higher energies and temperatures, based on their future electronic and nuclear motion calculations [118].

The calculation of the hot SO_2 line list has required a number of alterations to be made to the triatomic DVR3D code. The nature of the SO_2 system necessitates a large basis set to underlay all computations, and various procedures within the program structures have had to be optimised in order to deal with the large resulting wavefunction files. The future implementation of a refined PES should be fairly straightforward, although further optimisations may be necessary; for example if convergence tests re-

sult in the production of a higher density of computed bound states, particularly in the higher energy regions, then this would undoubtedly lead to an increase in the size of wavefunction files which in turn may result in storage issues, as well as longer computational times in dipole transition calculations. Nevertheless, the improvements to the DVR3D program suite in the current work have a number of implications beyond the case of the SO_2 molecule and could facilitate spectroscopic calculations for other heavy triatomic systems of interest, such as KCN [153, 247–251], TiO_2 [252, 253] and ClO_2 [254, 255].

7.1.2 The Hot SO_3 Line List

The SO_3 molecule has proven to be extremely difficult to work with experimentally and - owing to the formidable tasks associated with the current study - theoretically. The computation of a hot, *ab initio* SO_3 line list has tested the boundaries of some of the most sophisticated hardware and software implementations that are currently available. Even so, the final line list presented in this work has been the result of a compromise in the trade-off of accuracy and computational expense, and would certainly benefit from further studies in a number of ways.

The refinement procedure discussed in Chapter 6 has improved the position of many bands by a significant amount, and has managed to replicate very well their experimental counterparts. However a handful of theoretical band positions could benefit from further improvement, as alluded to in Table 6.2. This would require knowledge of a greater number of experimental bands in order to better refine the *ab initio* PES; a total of 119 experimentally derived rovibrational energy levels were used for the refinement in the current work for the hot SO_3 line list, with rotational excitation up to J=5. For example, compared to Ref. [239] for ammonia, which used 392 rovibrational terms in the fitting procedure, the number used in the current work remains relatively small and could be greatly improved by the availability of a larger number of reliable experimental measurements.

The complete lack of reliable absolute intensity data for SO_3 has made a robust validation of the *ab initio* DMS particularly difficult, and the only available means of comparison have been via relative intensity measurements by Maki *et al.* [99–105] and via semiquantitative cross section comparisons with a small number of data sets. In

particular, the experimental cross section data only became available after the production of the hot line list was started, and the semiquantitative analysis of cross section intensities for different temperatures suggests a temperature-independent defect in the calculated intensities. In other words, the discrepancies observed can probably be attributed to a defective DMS, as opposed to a temperature-dependent source, such as lack of partition function convergence. From initial comparisons there appears to be a constant factor defining the differences between experimental and theoretical results, which is unlike other cases in which *ab initio* DMSs have failed to accurately replicate observations [171, 245, 246]; this needs to be investigated. However, further studies into experimental cross sections would be required in order to reach a definitive conclusion, as there is a lack of observational data. In fact, determination of experimental absolute intensities would be the most desirable input, meaning comparisons could be made on a line-by-line basis. The current case for SO_3 provides a good example of the necessity of experimental and theoretical collaborations. In the meantime, the scaling of the theoretical intensity data produced in this work can facilitate analysis of experimental spectra for SO_3. A possible alternative avenue to explore in future would be in the systematic computational testing of the *ab initio* DMS, like those made by Lodi *et al.* for water [200]. The construction of a new *ab initio* DMS could also be a possibility, based on an improved theoretical calculation (such as the CCSD(T)/aug-cc-pV(Q+d)Z calculation in SO_2), to be used in conjunction with rovibrational wavefunctions that are still available from the calculations performed in this work. This has been discussed briefly in Chapter 6.

Finally, the coverage of the hot SO_3 line list is defined in terms of the wavenumber and temperature domains. The $E'' \leq 4000$ cm^{-1} and $J \leq 130$ cut-offs imposed by the calculation have resulted in the final line list being roughly 90% complete at $T = 773.15$ K. In fact, the $J \leq 130$ cut-off in particular means that there are a few energies below 4000 cm^{-1} that have still not been considered from higher J wavefunction calculations. Obtaining wavefunctions for $J > 130$ has proved to be extremely difficult using TROVE and the high performance computing systems available at the present time, however in theory these routines may be optimised in order to allow for computations to be successfully completed. Analysis of the partition function for $T = 773.15$ K suggests that line list completeness would be obtained by considering all transitions for (roughly)

$E'' \leq 6000$ cm^{-1}. Again, in theory this could be achieved using the wavefunction files currently available, however owing to impracticalities of computation cost this may be deemed as an unnecessary and expensive effort.

7.1.3 Modelling Astrophysical, Atmospheric & Industrial Environments

Chapter 1 discusses the astrophysical abundance of sulphur oxides. Missions dedicated to IR spectroscopic detection [61–63] can greatly benefit from the inclusion of both the SO_2 and SO_3 line lists in their analyses, particularly in the need for 'de-weeding' of SO_2 interference described in Ref. [64]. In situations where the molecule is of interest, the SO_2 line list can facilitate the theoretical modelling of solar system planetary atmospheres [13–15], and can be augmented by the inclusion of the SO_3 line list, where it may be considered as an equally major source of opacity [10, 11]. At higher temperatures the astrophysical modelling of SO_2 spectra may allow for a better understanding of its presence in stellar atmospheres. For example, in their study on the presence of SO_2 gas in oxygen-rich AGB stars [256], Yamamura *et al.* made use of the available HITRAN data in their spectral models, where they found modelling spectra at $T = 600$ K to give the best comparison between modelled and observed data for the ν_3 band. Figure 7.1 shows an example of such comparisons for Ref. [256] where they note that discrepancies in the wings of the ν_3 band are due to a lack of high energy transition data.

In a similar, more recent work, Adande *et al.* studied the sulphur chemistry in the envelope of the supergiant star VY Canis Majoris [24], where spectral models were based on CDMS data for the ground and $\nu_2 = 1$ vibrational state only, and on a simulated temperature of 400 K. Studies of this kind can benefit from the use of greater coverage of line positions provided by the hot SO_2 line list calculated in this work, which current spectroscopic databases may be lacking in.

The exciting implications of exoplanet characterisation efforts include the identification of 'Earth-like' planets [257]. The current state of both detection methods and spectroscopic technologies places a strong bias on the potentially observable objects, where the majority of detected exoplanets tend to be larger, gas giants close to their host star, which are more easily observed by current methods [257–259]. Nevertheless,

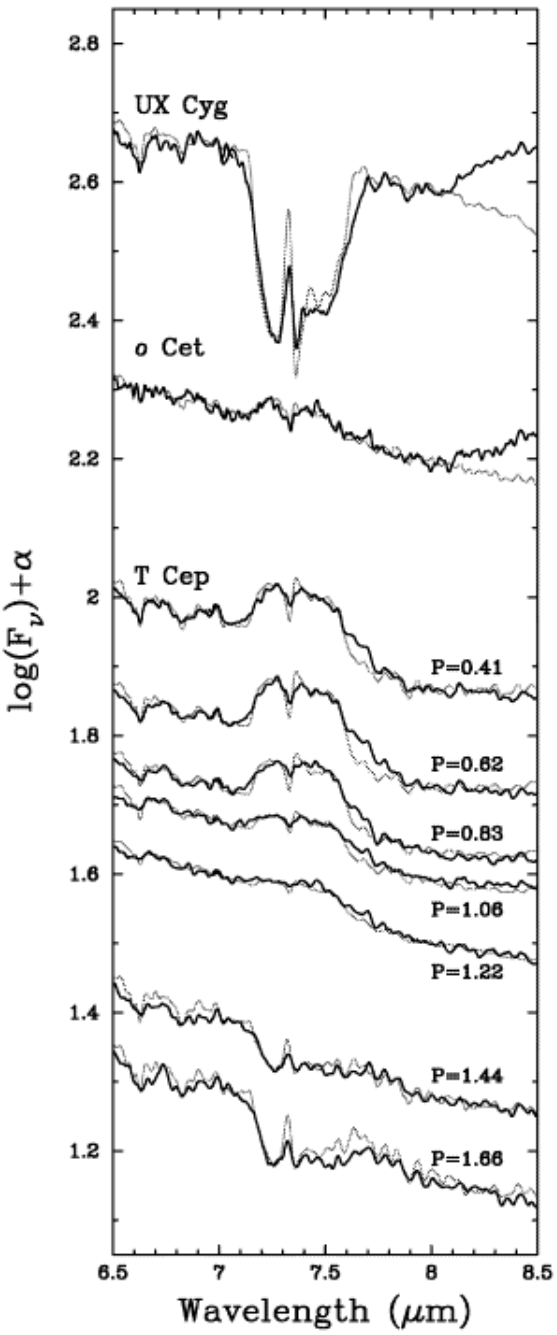

Figure 7.1: Figure taken from Ref. [256] showing observed spectra (thick line) compared against model spectra (thin line) using HITRAN data, for three oxygen-rich stars.

the detection of terrestrial-type planets is becoming more and more common [260,261] which, with ever improving technological advances, increases the potential for characterising those where SO_2 and SO_3 are likely to play a larger role [15, 260, 262, 263], as on Venus, Mars, Io and Earth.

Interest in the high-resolution and high-temperature spectroscopy of SO_2 and SO_3 within the context of the current project was originally stimulated by Alexander Fateev

and coworkers from the Technical University of Denmark (DTU). Their development of *in situ* measurement techniques for the purposes of gas temperature and composition within SCR units (see Chapter 1) requires high quality spectroscopic reference data, which have previously been limited in scope [244]. The importance of monitoring the various species concentrations within SCR units, particularly SO_2 and SO_3, is clear [5, 40, 41]. The high-temperature nature of the SO_2 and SO_3 line lists is likely to be immediately more important to such industrial applications. There have already been collaborative efforts between both the DTU and ExoMol groups involving the recording of hot ammonia spectra and line assignments via the BYTe line list [244], in order to better understand the detection of ammonia within the SCR environment. Both the SO_2 and SO_3 line lists may be used in a similar fashion in conjunction with experimental observations at elevated temperatures. In particular it is hoped that they may provide further insight into SO_3 abundances, where laboratory measurements are prone to difficulty due to losses of this highly reactive species within gas flow cells [42]. The high-temperature line lists computed in the current work can also be used to supplement other studies involved in the remote sensing of sulphur dioxide in other exhaust gases [39, 116].

7.1.4 SO_2 & SO_3 Line Widths

Spectral lines observed in nature are usually subject to broadening due to temperature and pressure effects associated with their environment, and do not exhibit an infinitely thin line width as given by the computed line lists in this work. Spectral features will always be subject to a natural broadening depending on the lifetime of an excited state. Doppler broadening, which is associated with the thermal motion of particles, can give an indication as to the temperature of the objects under observation, manifesting in an increase of the spectral line width. Mathematically, the shape exhibited by thermal broadening can be described by a Gaussian line shape profile [264]. The line shape can also be affected by the pressure of the environment, where the constant collisions between molecules leads to a distortion of energy levels and their lifetimes. This type of broadening is best described by a Lorentz profile. In reality temperature and pressure are not mutually exclusive, and therefore must both be accounted for in order to more accurately simulate molecular absorption cross sections for environments of interest;

the combination of Gaussian and Lorentz profiles leads to the so-called 'Voigt' profile [244, 265].

The cross section data produced in this work for both SO_2 and SO_3 in Chapters 3 and 6, respectively, are constructed based on a Gaussian line shape function used to convolve transition lines from the line lists, i.e. they only consider Doppler broadening. To simulate pressure effects it is necessary to implement a Voigt profile which is difficult to create, and requires the use of pressure broadening parameters to achieve the appropriate shape. The two important parameters of interest are the Lorentz half-width, γ, and the temperature dependent exponent, n, which scales γ recorded at a reference temperature to the temperature of interest [265].

There have been a number of efforts to obtain accurate broadening parameters for SO_2, both experimentally and theoretically. Sumpf *et al.* investigated the effects of self-broadening [266–268], air-broadening [269] and Nitrogen-broadening [270] for the ν_1, ν_2, ν_3 and $\nu_1 + \nu_3$ bands, as well as broadening by noble gases [271–273]. Tasinato *et al.* also reported self-broadening parameters in the 9.2 μm spectral region [274]. Broadening by H_2, N_2, O_2 and He has also been studied by various authors [275–277], and measurement of a single measured line has provided some insight into pressure broadening due to CO_2 [278]. A number of broadening parameters obtained from these studies have also been incorporated into the HITRAN data set for this molecule [265]. Conversely, there have been no such studies published for the SO_3 molecule.

The use of Voigt profiling based on line list data has already been carried out in the previously mentioned collaboration between the DTU and ExoMol groups in order to simulate and assign hot ammonia spectra using the BYTe line list [244], making use of Lorentz half-width data available from HITRAN. An identical approach with SO_2 and SO_3 in conjunction with available broadening parameter data should enable for an improved description of cross section characteristics at high temperature. The available data for pressure broadening for SO_2 should make it possible to begin investigating this, however corresponding data for the SO_3 would also be desirable. In the meantime Doppler broadened cross sections for SO_2 and SO_3 created in this work may be used for modelling environments in hot conditions.

7.1.5 Cluster State Analysis for SO_2

The theoretical rotational energy cluster formation observed for SO_3 in this work suggests that a similar phenomenon in SO_2 may be possible. As previously stated, rotational energy clustering is seen mainly in XY_N type molecules with a heavy central atom X, with Y being H or D, and X-Y_i bonds being near-orthogonal, i.e. local mode molecules [167, 221], as seen in H_2S, H_2Se, H_2Te, and PH_2 [167, 214, 221–223]. SO_2 shares the similar triatomic nature of these XY_N molecules, and also contains a relatively heavy central X atom, though can not strictly be defined by local mode behaviour due to the 119° angle subtended by the S-O bonds. However, the same is true for SO_3 which has theoretically been shown to exhibit this clustering phenomenon, therefore suggesting that this behaviour is worth investigating in SO_2. The rotational energy clustering is observed for high rotational excitation in SO_3, which required a reduced basis set in the TROVE calculations to produce the high energy energy states in a reasonable time. Rotational excitation was considered as high as $J = 250$ in the vibrational ground state, and clustering behaviour exhibited from J=125 onwards; the DVR3D calculations for the SO_2 line list considered rovibrational energy calculations up to $J = 160$. It is therefore possible that clustering phenomena may already be observed within the already available calculations. The analysis is not as straightforward as with SO_3 and TROVE as the DVR3D calculations do not readily provide vibrational quantum numbers, making it difficult to separate out individual rotational levels within vibrational states. The ever-increasing density of states with higher rotational excitation would also make this process very difficult, therefore the DVR3D code would require some changes in order to specify the vibrational contribution for each rovibrational wavefunction. Alternatively, analysis may be performed using the TROVE program. In Chapter 3 the use of TROVE with the Ames-1 PES was discussed briefly, where it successfully reproduced DVR3D-computed energies for J=0 energy levels. It may therefore be possible to implement the Ames-1B PES used in the calculation of the hot SO_2 line list in order to produce high-J energies labelled by vibrational quanta, making the assessment of clustering phenomena more straightforward.

7.2 Conclusion

The SO_2 and SO_3 hot line lists should help provide an extensive insight into the characteristics of these molecules for a large range of temperatures and across a wide spectral window, which in turn will provide experimental spectroscopists with a means to further understand the environments in which they have significant presence. Both line lists contribute to an extensive improvement in spectral data coverage as part of the ExoMol project, and the collaborative efforts between the ExoMol group and DTU represents a successful marriage of academic and industrial disciplines.

New avenues of investigation may opened up based on the results of this project, from the assessment of *ab initio* quantum chemistry methods, and the further improvement of efficient computational algorithms, to the improved characterisation of atmospheric behaviour, within industrial, atmospheric and extraterrestrial environments.

The work presented in this thesis has been the result of over three years' effort, requiring a detailed knowledge of a variety of fundamental physical concepts and their implementation into sophisticated computer codes optimised for efficient use on high performance computing systems, followed by thorough analyses of the huge quantity of produced data. The results presented are due to a combination of both independent research and collaboration with scientists from various fields, including spectroscopy, quantum chemistry, chemical engineering and planetary atmospheres.

Appendix A

SO_3 Refinement Energies

Tables A.1 to A.5 contain the remaining comparisons for the rovibrational energy levels used in the refinement procedure for the UYT2 SO_3 PES, discussed in Chapter 6.

Table A.1: Obs. - Calc. residuals for the $J = 0$ energy levels used in the refinement procedure. All values are in cm^{-1}.

State	K	Obs. [104]	UYT2	Obs. - Calc.
$3\nu_4^{(l_4=1)}$	0	1591.097	1591.040	0.057
$2\nu_3^{(l_3=2)}$	0	2766.405	2766.384	0.022

⋆The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

Table A.2: Obs. - Calc. residuals for the $J = 1$ energy levels used in the refinement procedure. All values are in cm^{-1}.

State	K	Obs. [104]	UYT2	Obs. - Calc.
ν_2	0	498.257	498.246	0.010
$\nu_3^{(l_3=1)}$	1	1391.874	1391.863	0.011
$2\nu_2 + \nu_4^{(l_4=1)}$	1	1524.897	1524.909	-0.012
$\nu_1 + \nu_2$	0	1565.774	1565.757	0.017
$2\nu_3^{(l_3=2)}$	1	2778.723	2778.726	-0.003
$3\nu_3^{(l_3=1)}$	1	4136.740	4136.731	0.009

⋆The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

Table A.3: Obs. - Calc. residuals for the $J = 2$ energy levels used in the refinement procedure. All values are in cm^{-1}.

State	K	Obs. [104]	UYT2	Obs. - Calc.
ν_0	0	2.091	2.092	0.000
$\nu_4^{(l_4=1)}$	2	531.138	531.137	0.001
	1	532.174	532.183	-0.009
$\nu_2 + \nu_4^{(l_4=1)}$	2	1028.938	1028.903	0.035
$\nu_3^{(l_3=1)}$	2	1393.243	1393.242	0.001
	1	1393.264	1393.253	0.011
$2\nu_2 + \nu_4^{(l_4=1)}$	2	1525.220	1525.223	-0.003
	1	1526.254	1526.267	-0.013
$3\nu_4^{(l_4=1)}$	2	1590.046	1590.028	0.018
$^{\star}(\nu_3 + \nu_4)^{(L=2)}$	2	1919.085	1919.094	-0.009
	1	1919.590	1919.597	-0.007
$2\nu_3^{(l_3=2)}$	2	2778.591	2778.575	0.016
	1	2780.108	2780.111	-0.004
$3\nu_3^{(l_3=1)}$	2	4138.095	4138.097	-0.002

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

Table A.4: Obs. - Calc. residuals for the $J = 3$ energy levels used in the refinement procedure. All values are in cm^{-1}.

State	K	Obs. [104]	UYT2	Obs. - Calc.
ν_0	3	2.612	2.612	0.000
ν_2	3	500.168	500.158	0.010
	0	501.701	501.692	0.009
$\nu_4^{(l_4=1)}$	2	533.243	533.242	0.001
	1	534.320	534.329	-0.009
$2\nu_2$	3	996.222	996.273	-0.051
$\nu_2 + \nu_4^{(l_4=1)}$	2	1031.016	1030.981	0.035
	1	1032.041	1032.016	0.025
$2\nu_4^{(l_4=2)}$	2	1064.664	1064.671	-0.007
$\nu_3^{(l_3=1)}$	2	1395.327	1395.327	0.000
	1	1395.347	1395.337	0.010
$2\nu_2 + \nu_4^{(l_4=1)}$	2	1527.275	1527.279	-0.004
	1	1528.338	1528.351	-0.013
$\nu_1 + \nu_2$	3	1567.682	1567.665	0.017
	0	1569.212	1569.196	0.016
$\nu_1 + \nu_4^{(l_4=1)}$	2	1599.089	1599.086	0.004
	1	1600.178	1600.182	-0.004
$^\star(\nu_3 + \nu_4)^{(L=2)}$	2	1921.185	1921.214	-0.029
	1	1921.688	1921.695	-0.007
$2\nu_3^{(l_3=2)}$	2	2780.669	2780.653	0.016
	1	2782.186	2782.190	-0.004
$3\nu_3^{(l_3=1)}$	2	4140.167	4140.114	0.053
	1	4140.191	4140.129	0.062

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

Table A.5: Obs. - Calc. residuals for the $J = 4$ energy levels used in the refinement procedure. All values are in cm^{-1}.

State	K	Obs. [104]	UYT2	Obs. - Calc.
ν_0	3	5.400	5.401	-0.001
	0	6.971	6.972	-0.001
ν_2	3	502.923	502.915	0.009
$\nu_4^{(l_4=1)}$	4	534.963	534.982	-0.019
	2	536.051	536.050	0.001
	1	537.056	537.066	-0.010
$2\nu_2$	3	998.948	999.001	-0.052
	0	1000.448	1000.501	-0.054
$\nu_2 + \nu_4^{(l_4=1)}$	4	1032.778	1032.762	0.017
	2	1033.787	1033.753	0.034
	1	1034.871	1034.846	0.025
$2\nu_4^{(l_4=0)}$	3	1064.926	1064.919	0.006
	0	1066.532	1066.525	0.006
$2\nu_4^{(l_4=2)}$	1	1066.994	1066.982	0.012
	2	1067.497	1067.504	-0.007
$\nu_3^{(l_3=1)}$	4	1395.016	1394.996	0.020
	2	1398.106	1398.107	0.000
	1	1398.129	1398.119	0.010
$2\nu_2 + \nu_4^{(l_4=1)}$	4	1529.073	1529.094	-0.021
	2	1530.015	1530.021	-0.006
	1	1531.028	1531.043	-0.015
$\nu_1 + \nu_2$	3	1570.432	1570.417	0.015
$3\nu_4^{(l_4=1)}$	3	1593.868	1593.870	-0.002
	2	1595.007	1594.990	0.017
$\nu_1 + \nu_4^{(l_4=1)}$	4	1600.787	1600.805	-0.018
	2	1601.907	1601.903	0.004
	1	1602.892	1602.897	-0.005
$^\star(\nu_3 + \nu_4)^{(L=2)}$	4	1921.813	1921.821	-0.008
	2	1923.981	1923.990	-0.009
	1	1924.485	1924.493	-0.008
$2\nu_3^{(l_3=2)}$	4	2783.349	2783.373	-0.024
	2	2783.439	2783.424	0.015
	1	2784.955	2784.960	-0.005
$3\nu_3^{(l_3=1)}$	4	4139.865	4139.793	0.072
	2	4142.927	4142.876	0.051
	1	4142.955	4142.894	0.061

*The value L is given by $L = |l_3 + l_4|$ as per Ref. [105].

Bibliography

[1] R. E. Stoiber and A. Jepsen. Sulfur dioxide contributions to the atmosphere by volcanoes. *Science*, 182:577–578, 1973.

[2] J. P. Michaud, D. Krupitsky, J. S. Grove, and B. S. Anderson. Volcano related atmospheric toxicants in hilo and hawaii volcanoes national park: Implications for human health. *Neurotoxicology*, 26:555–563, 2005.

[3] A. S. Khayat, G. L. Villanueva, M. J. Mumma, and A. T. Tokunaga. A search for SO2, H2S and SO above Tharsis and Syrtis volcanic districts on Mars using ground-based high-resolution submillimeter spectroscopy. *Icarus*, 253:130–141, 2015.

[4] A. R. Whitehill, C. Xie, X. Hu, D. Xie, H. Guo, and S. Ono. Vibronic origin of sulfur mass-independent isotope effect in photoexcitation of SO2 and the implications to the early earth's atmosphere. *PNAS*, 110:17697–17702, 2013.

[5] M. Vahedpour, M. Goodarzi, N. Hajari, and F. Nazari. Theoretical study on the atmospheric formation of sulfur trioxide as the primary agent for acid rain. *Structural Chemistry*, 22:817–822, 2011.

[6] C. E. Kolb, J. T. Jayne, and D. R. Worsnop. Gas phase reaction of sulphur trioxide with water. *J. Am. Chem. Soc.*, 116:10314–10315, 1994.

[7] R. K. Srivastava, C. A. Miller, C. Erickson, and R. Jambheka. Emissions of sulfur trioxide from coal-fired power plant. *Journal of the Air and Waste Management Association*, 54(6):750–762, 2004.

[8] T. Hieta and M. Merimaa. Simultaneous detection of SO2, SO3 and H2O using QCL spectrometer for combustion applications. *Appl. Phys. B*, 117:847–854, 2014.

[9] C. Y. Na, L. W. Esposito, and T. E. Skinner. International ultraviolet explorer observation of Venus SO2 and SO. *Journal of Geophysical Research*, 95:7485–7491, 1990.

[10] X. Zhang, M. Liang, F. Montmessin, J. Bertaux, C. Parkinson, and Y. L. Yung. Photolysis of sulphuric acid as the source of sulphur oxides in the mesosphere of Venus. *Nat. Geosci.*, pages 834–837, 2010.

[11] X. Zhang, M. C. Liang, F. P. Mills, D. A. Belyaev, and Y. L. Yung. Sulfur chemistry in the middle atmosphere of Venus. *Icarus*, 217:714–739, 2012.

[12] R. A. Craig, R. T. Reynolds, B. Ragent, G. C. Carle, F. Woeller, and J. B. Pollack. Sulfur trioxide in the lower atmosphere of Venus? *Icarus*, 53:1–9, 1983.

[13] D. A. Belyaev, F. Montmessin, J.-L. Bertaux, A. Mahieux, A. A. Fedorova, O. I. Korablev, E. Marcq, Y. L. Yung, and X. Zhang. Vertical profiling of SO2 and SO above Venus clouds by SPICAV/SOIR solar occultations. *Icarus*, 217:740–751, 2012.

[14] D. Belyaev, O. Korablev, A. Fedorova, J.-L. Bertaux, A.-C. Vandaele, F. Montmessin, A. Mahieux, V. Wilquet, and R. Drummond. First observations of SO2 above Venus clouds by means of Solar Occultation in the Infrared. *Journal of Geophysical Research*, 113, 2008.

[15] G. Arney, V. Meadows, D. Crisp, S. J. Schmidt, J. Bailey, and T. Robinson. Spatially resolved measurements of H2O, HCl, CO, OCS, SO2, cloud opacity, and acid concentration in the Venus near-infrared spectral windows. *Journal of Geophysical Research: Planets*, 119:1860–1891, 2014.

[16] J. Pearl, R. Hanel, V. Kunde, W. Maguire, K. Fox, S. Gupta, C. Ponnamperuma, and F. Raulin. Identification of gaseous SO2 and new upper limits for other gases on Io. *Nature*, 280:755 – 758, 1979.

[17] R. M. Nelson, A. L. Lane, D. L. Matson, F. P. Fanale, D. B. Nash, , and T. V. Johnson. Longitudinal distribution of SO2 on Io. *Science*, 210:784, 1980.

[18] M. H. Moore. Studies of proton-irradiated SO2 at low temperatures: Implications for Io. *Icarus*, 59:114–128, 1984.

[19] G. E. Ballester, M. A. McGrath, D. F. Strobel, X. Zhu, P. D. Feldman, and H. W. Moos. Detection of the SO2 Atmosphere on Io with the Hubble Space Telescope. *Icarus*, 111:2–17, 1994.

[20] C. Visscher, K. Lodders, and B. Fegley Jr. Atmospheric chemistry in giant planets, brown dwarfs, and low-mass dwarf stars. II. Sulfur and Phosphorus. *Astrophys. J.*, 648:1181, 2006.

[21] I. Yamamura, T. de Jong, T. Osaka, J. Cami, and L. B. F. M. Waters. Detection of warm SO2 gas in oxygen-rich AGB stars. *Astron. Astrophys.*, 341:L9–L12, 1999.

[22] F. F. S. van der Tak, A. M. S. Boonman, R. Braakman, , and E. F. van Dishoeck. Sulfur chemistry in the envelopes of massive young stars. *Astron. Astrophys.*, 412:133, 2003.

[23] L. M. Ziurys. The chemistry in circumstellar envelopes of evolved stars: Following the origin of the elements to the origin of life. *Proc. Natl. Acad. Sci. U.S.A.*, 103:12274, 2006.

[24] G. R. Adande, J. L. Edwards, and L. M. Ziurys. Sulfur chemistry in the envelope of VY Canis Majoris: Detailed Analysis of SO and SO2 emission. *Astrophysical Journal*, 778:1–11, 2013.

[25] L. M. Chernin, C. R. Masson, and G. A. Fuller. High-velocity sulfur monoxide emission from protostellar outflows. *Astrophysical Journal, Part 1*, 436:741–748, 1994.

[26] E. Klisch, P. Schilke, S. P. Belov, , and G. Winnewisser. $^{33}SO_2$: Interstellar identification and laboratory measurements. *J. Mol. Spectrosc.*, 186:314, 1997.

[27] P. Schilke, D. J. Benford, T. R. Hunter, D. C. Lis, , and T. G. Philips. A line survery of Orion-KL from 607 to 725 GHz. *Astrophys. J., Suppl. Ser.*, 132:281, 2001.

[28] N. R. Crockett, E. A. Bergin, S. Wang, D. C. Lis, and T. A. Bell et al. Herschel observations of extra-ordinary sources (HEXOS): The terahertz spectrum of Orion KL seen at high spectral resolution. *Astron. Astrophys.*, 521:L21, 2010.

[29] A. Belloche, H. S. P. Mller, K. M. Menten, P. Schilke, , and C. Comito. Complex organic molecules in the interstellar medium: IRAM 30 m line survey of Sagittarius B2(N) and (M). *Astron. Astropys.*, 559:A47, 2013.

[30] C. A. Gottlieb and J. A. Ball. Interstellar sulfur monoxide. *Astrophysical Journal, Part 1*, 184:L59, 1973.

[31] C. A. Gottlieb, E. W. Gottlieb, M. M. Litvak, J. A. Ball, and H. Penfield. Observations of interstellar sulfur monoxide. *Astrophysical Journal, Part 1*, 219:77–94, 1978.

[32] S. Martin, R. Mauersberger, J. Martin-Pintado, S. Garcia-Burillo, and C. Henkel. First detections of extragalactic SO2, NS and NO. *Astron. Astrophys.*, 411:L465, 2003.

[33] S. Martin, J. Martin-Pintado, R. Mauersberger, C. Henkel, and S. Garcia-Burillo. Sulfur chemistry and isotopic ratios in the starburst galaxy NGC 253. *Astrophys. J.*, 620:210, 2005.

[34] S. J. Petuchowski and C. L. Bennett. SO in starburst galaxies. *Astrophysical Journal, Part 1*, 391:137–140, 1992.

[35] X. Huang, D. W. Schwenke, and T. J. Lee. Highly accurate potential energy surface, dipole moment surface, rovibrational energy levels, and infrared line list for $^{32}S^{16}O_2$ up to 8000 cm^{-1}. *J. Chem. Phys.*, 140:114311, 2014.

[36] M. Kama, A. Lopez-Sepulcre, C. Dominik, C. Ceccarelli, A. Fuente, and E. Caux et al. The Herschel/HIFI spectral survey of OMC-2 FIR 4 (CHESS): An overview of the 480 to 1902 GHz range. *Astron. Astrophys.*, 556:A57, 2013.

[37] C. L. Huang, S. S. Ju, I. C. Chen, A. J. Merer, C. K. Ni, and A. H. Kung. High-resolution spectroscopy of jet-cooled $^{32}SO_2$ and $^{34}SO_2$: The $\tilde{a}^3B_1$-$\tilde{X}^1A_1$, 2_0^1 and 1_0^1 bands. *J. Mol. Spectrosc.*, 203:151–157, 2000.

[38] C. L. Huang, I. C. Chen, A. J. Merer, C. K. Ni, and A. H. Kung. Spectra of jet-cooled $^{32}SO_2$ and $^{34}SO_2$ in systems $\tilde{a}^3B_1$ and $\tilde{b}^3A_2$-$\tilde{x}^1A_1$: Rotational structure of perturbed $\tilde{b}^3A_2$. *J. Chem. Phys.*, 114:1187–1193, 2001.

[39] O. K. Voitsekhovskaya, D. E. Kashirskii, and O. V. Egorov. Spectroscopic support of laser remote sensing of the sulfur dioxide gas in the jet of engine exhaust gases. *Russian Phys. J.*, 56:473–482, 2013.

[40] W. T. Rawlins, J. M. Hensley, D. M. Sonnenfroh, D. B. Oakes, and M. G. Allen. Quantum cascade laser sensor for SO2 and SO3 for application to combustor exhaust streams. *Appl. Optics*, 44:6635–6643, 2005.

[41] D. Fleig, E. Vainio K., Andersson, A. Brink, F. Johnsson, and M. Hupa. Evaluation of SO3 Measurement Techniques in Air and Oxy-Fuel Combustion. *Energy Fuels*, 26:5537–5549, 2012.

[42] H. Grosch, A. Fateev, K. L. Nielsen, and S. Clausen. Hot gas flow cell for optical measurements on reactive gases. *J. Quant. Spectrosc. Radiat. Transf.*, 130:392–399, 2013.

[43] H. Grosch, A. Fateev, and S. Clausen. UV absorption cross-sections of selected sulfur-containing compounds at temperatures up to 500C. *JQSRT*, 154:28–34, 2014.

[44] E. S. Barker. Detection of SO2 in the UV Spectrum of Venus. *JGR*, 6:117–120, 1979.

[45] S. A. Carna, A. J. Kruegera, S. Arellanob, N. A. Krotkovc, and K. Yang. Daily monitoring of ecuadorian volcanic degassing from space. *Journal of Volcanology and Geothermal Research*, 176:141–150, 2008.

[46] A. J. Krueger. Sighting of El Chichon Sulfur Dioxide Clouds with the Nimbus 7 Total Ozone Mapping Spectrometer. *Science*, 220:1377–1379, 1983.

[47] A. J. Krueger, L. S. Walter, P. K. Bhartia, C. C. Schnetzler, N. A. Krotkov, I. Sprod, and G. J. S. Bluth. Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments. *Journal of Geophysical Research: Atmospheres*, 100:14057–14076, 1995.

[48] M. Eisinger and J. P. Burrows. Tropospheric sulfur dioxide observed by the ERS-2 GOME instrument. *Geophysical Research Letters*, 25:4177–4180, 1998.

[49] J. Rufus, G. Stark, P. L. Smith, J. C. Pickering, and A. P. Thorne. High-resolution photoabsorption cross section measurements of SO2, 2: 220 to 325 nm at 295 K. *J. Geographical Research.*, 108(E2):5011, 2003.

[50] J. S. Baskin, F. Aladel, and A. Hamdan. Unexpectedly rich vibronic structure in supersonic jet spectra of sulfur-dioxide between 360 and 308 nm. *J. Chem. Phys.*, 200:181–199, 1995.

[51] J. A. Joens. Alternative assignments for the vibrational structure of the $^3B_1$-X^1A_1 band system of SO2. *Chem. Phys. Lett.*, 261:659–664, 1996.

[52] A. C. Vandaele, C. Hermans, and S. Fally. Fourier transform measurements of SO2 absorption cross sections: II. Temperature dependence in the 29000-44000 cm^{-1} (227-345 nm) region. *J. Quant. Spectrosc. Radiat. Transf.*, 110:2115–2126, 2009.

[53] D. Q. Xie, H. Guo, O. Bludsky, and P. Nachtigall. Absorption and resonance emission spectra of SO2($\tilde{X}^1A_1/\tilde{C}^1B_2$) calculated from ab initio potential energy and transition dipole moment surfaces. *Chem. Phys. Lett.*, 329:503–510, 2000.

[54] H. Ran, D. Xie, and H. Guo. Theoretical studies of $\tilde{C}^1B_2$ absorption spectra of SO2 isotoporners. *Chem. Phys. Lett.*, 439:280–283, 2007.

[55] I. Tokue and S. Nanbu. Theoretical studies of absorption cross sections for the $\tilde{C}^1B_2$-$\tilde{X}^1A_1$ system of sulfur dioxide and isotope effects. *J. Chem. Phys.*, 132:024301, 2010.

[56] C. Leveque, A. Komainda, R. Taeb, and H. Koppel. Ab initio quantum study of the photodynamics and absorption spectrum for the coupled 1^1 A_2 and 1^1B_1 states of SO2. *J. Chem. Phys.*, 138, 2013.

[57] C. Xie, X. Hu, L. Zhou, D. Xie, and H. Guo. Ab initio determination of potential energy surfaces for the first two UV absorption bands of SO2. *J. Chem. Phys.*, 139:014305, 2013.

[58] E. Fajans and C. F. Goodeve. The absorption spectrum of sulphur trioxide. *Trans. Faraday Soc.*, 32:511–514, 1936.

[59] B. Leroy, G. Le Bras, and P. Rigaud. Ultraviolet absorption spectra of sulfur compounds of aeronomical interest (carbon oxysulfide, carbon disulfide, sulfur trioxide). *Annales de Geophysique*, 37:297–302, 1981.

[60] J. B. Burkholder and S. McKeen. UV absorption cross sections for SO3. *Geophysical Research Letters*, 24:3201–3204, 1997.

[61] G. L. Pilbratt, J. R. Riedinger, T. Passvogel, G. Crone, D. Doyle, U. Gageur, A. M. Heras, C. Jewell, L. Metcalfe, S. Ott, and M. Schmidt. Herschel Space Observatory: An ESA facility for far-infrared and submillimetre astronomy. *A& A*, 518:L1, 2010.

[62] M. W. Werner, T. L. Roellig, F. J. Low, G. H. Rieke, M. Rieke, W. F. Hoffmann, E. Young, J. R. Houck, B. Brandl, and G. G. Fazio. The spitzer space telescope mission. *The American Astronomical Society*, 154:1–9, 2004.

[63] E. T. Young, E. E. Becklin, P. M. Marcum, T. L. Roellig, J. M. De Buizer, T. L. Herter, R. Gusten, E. W. Dunham, P. Temi, and B.G. Andersson. Early science with SOFIA, the stratospheric observatory for infrared astronomy. *The American Astronomical Society*, 749:L17, 2012.

[64] T. Bergin, F. De. Lucia, P. Goldsmith, H. Hasan, E. Herbst, G. Melnick, and J. Pearson. Report from the workshop on laboratory spectroscopy in support of HERSCHEL, SOFIA, and ALMA. *Pasadena, CA*, 2006.

[65] G. F. Crable and W. V. Smith. The structure and dipole moment of SO2 from microwave spectra. *J. Chem. Phys.*, 19:502, 1951.

[66] R. J. Corice, K. Fox, and G. D. T. Tejwani. Experimental and theoretical studies of the fundamental bands of sulfur dioxide. *J. Chem. Phys.*, 58:265–270, 1973.

[67] D. Patel, D. Margolese, and T. R. Dyke. Electric-dipole moment of SO2 in ground and excited vibrational-states. *J. Chem. Phys.*, 70:2740–2747, 1979.

[68] J.-M. Flaud, A. Perrin, L. M. Salah, W. J. Lafferty, and G. Guelachvili. A reanalysis of the (010), (020), (100), and (001) rotational levels of $^{32}S^{16}O_2$. *Journal of Molecular Spectroscopy*, 160:272–278, 1993.

[69] P. M. Chu, S. J. Wetzel, W. J. Lafferty, A. Perrin, J. M. Flaud, P. Arcas, and G. Guelachvili. Line intensities for the 8 μm bands of SO2. *J. Mol. Spectrosc.*, 189:55–63, 1998.

[70] W. J. Lafferty, A. S. Pine, G. Hilpert, R. L. Sams, and J.-M. Flaud. The $\nu_1 + \nu_3$ and $2\nu_1 + \nu_3$ band systems of SO2: line positions and intensities. *Journal of Molecular Spectroscopy*, 176:280–286, 1996.

[71] W.J. Lafferty, G.T. Fraser, A.S. Pine, J.-M. Flaud, C. Camy-Peyret, V. Dana, J.-Y. Mandin, A. Barbe, J.-J. Plateaux, and S. Bouazza. The 33 band of 32s16o2: line positions and intensities. *Journal of Molecular Spectroscopy*, 154:51–60, 1992.

[72] G. D. T. Tejwani, K. Fox, and R. J. Corice. Dipole-moment derivatives for SO2. *Chem. Phys. Lett.*, 18:365–368, 1973.

[73] F. Vonmoers, K. Schmalzigaug, A. Schmoldt, S. Heitz, and A. Hese. Excited-state dipole-moments of SO2. *J. Chem. Phys.*, 92:6391–6396, 1990.

[74] C. C. Zen, I. C. Chen, Y. P. Lee, and A. J. Merer. Laser-induced phosphorescence of SO2 in solid neon: Direct observation of the $\tilde{b}^3A_2$ state in the $^{16}OS^{18}O$ molecule. *J. Phys. Chem. A*, 104:771–776, 2000.

[75] L. S. Rothman, I. E. Gordon, Y. Babikov, A. Barbe, D. Chris Benner, P. F. Bernath, M. Birk, L. Bizzocchi, V. Boudon, L. R. Brown, A. Campargue,

K. Chance, E. A. Cohen, L. H. Coudert, V. M. Devi, B. J. Drouin, A. Fayt, J.-M. Flaud, R. R. Gamache, J. J. Harrison, J.-M. Hartmann, C. Hill, J. T. Hodges, D. Jacquemart, A. Jolly, J. Lamouroux, R. J. Le Roy, G. Li, D. A. Long, O. M. Lyulin, C. J. Mackie, S. T. Massie, S. Mikhailenko, H. S. P. Müller, O. V. Naumenko, A. V. Nikitin, J. Orphal, V. Perevalov, A. Perrin, E. R. Polovtseva, C. Richard, M. A. H. Smith, E. Starikova, K. Sung, S. Tashkun, J. Tennyson, G. C. Toon, Vl. G. Tyuterev, and G. Wagner. The *HITRAN* 2012 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 130:4 – 50, 2013.

[76] H. S. P. Müller, F. Schlöder, J. Stutzki, and G. Winnewisser. The Cologne database formolecularspectroscopy, CDMS: a useful tool for astronomers and-spectroscopists. *J. Molecular Structure*, 742:215–227, 2005.

[77] W. J. Lafferty, A. S. Pine, J.-M. Flaud, and C. Camy-Peyret. The $2\nu_3$ band of $^{32}S^{16}O_2$: line positions and intensities. *Journal of Molecular Spectroscopy*, 157:499–511, 1993.

[78] J.-M. Flaud and W. J. Lafferty. $^{32}s^{16}o_2$: a refined analysis of the $3\nu_3$ band and determination of equilibrium rotational constants. *Journal of Molecular Spectroscopy*, 161:396–402, 1993.

[79] J. Henningsen, A. Barbe, and M.-R. De Backer-Barilly. Revised molecular parameters for $^{32}SO_2$ and $^{34}SO_2$ from high resolution study of the infrared spectrum in the 7-8 μm wavelength region. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109:2491–2510, 2008.

[80] J. R. Spencer, E. Lellouch, M. J. Richter, M. A. Lopez-Valverde, K. L. Jessup, T. K. Greathouse, and J.-M. Flaud. Mid-infrared detection of large longitudinal asymmetries in Io's SO2 atmosphere. *Icarus*, 176:283–304, 2005.

[81] O. N. Ulenikov, E. S. Bekhtereva, S. Alanko, V.-M. Horneman, O. V. Gromova, and C. Leroy. On the high resolution spectroscopy and intramolecular potential function of SO2. *J. Mol. Spectrosc.*, 257:137–156, 2009.

[82] O. N. Ulenikov, O. V. Gromova, E. S. Bekhtereva, I. B. Bolotova, C. Leroy, V.-M. Horneman, and S. Alanko. High resolution study of the $\nu_1 + 2\nu_2 - \nu_2$ and

$2\nu_2+\nu_3-\nu_2$ hot bands and ro-vibrational re-analysis of the $\nu_1+\nu_2/\nu_2+\nu_3/3\nu_2$ polyad of the $^{32}SO_2$ molecule. *J. Quant. Spectrosc. Radiat. Transfer*, 112:486–512, 2011.

[83] O. N. Ulenikov, O. V. Gromova, E. S. Bekhtereva, I. B. Bolotova, I. A. Konov, V.-M. Horneman, and C. Leroy. High resolution analysis of the SO2 spectrum in the 2600View the MathML source region: $2\nu_3$, $\nu_2+2\nu_3-\nu_2$ and $2\nu_1+\nu_2$ bands. *J. Quant. Spectrosc. Radiat. Transfer*, 113:500–517, 2012.

[84] O. N. Ulenikov, E. S. Bekhtereva, V.-M. Horneman, S. Alanko, , and O. V. Gromova. High resolution study of the $3\nu_1$ band of SO2. *J. Mol. Spectrosc.*, 255:111–121, 2009.

[85] O. N. Ulenikov, E. S. Bekhtereva, O. V. Gromova, S. Alanko, V.-M. Horneman, and C. Leroy. Analysis of highly excited hot bands in the SO2 molecule: $\nu_2+3\nu_3-\nu_2$ and $2\nu_1+\nu_2+\nu_3-\nu_2$. *Mol. Phys.*, 108:1253–1261, 2010.

[86] O. N. Ulenikov, G. A. Onopenko, O. V. Gromova, E. S. Bekhtereva, and V.-M. Horneman. Re-analysis of the (100), (001), and (020) rotational structure of SO2 on the basis of high resolution FTIR spectra. *J. Quant. Spectrosc. Radiat. Transfer*, 130:220–232, 2013.

[87] Y. Endo, S. O. Danielache, Y. Ueno, S. Hattori, M. S. Johnson, N. Yoshida, and H. G. Kjaergaard. Photoabsorption cross-section measurements of S-32, S-33, S-34, and S-36 sulfur dioxide from 190 to 220 nm. *Journal of Geophysical Research*, 120:2546–2557, 2015.

[88] G. Stark, P. L. Smith, J. Rufus, A. P. Thorne, J. C. Pickering, and G. Cox. High-resolution photoabsorption cross-section measurements of SO2 at 295 K between 198 and 220 nm. *Journal of Geophysical Research*, 104:6585–6590, 1999.

[89] J. Rufus, G. Stark, A. P. Thorne, J. C. Pickering, R. J. Blackwell-Whitehead, D. Blackie, and Peter L. Smith. High-resolution photoabsorption cross-section measurements of SO2 at 160 K between 199 and 220 nm. *J. Geophysical Research*, 114:E06003, 2009.

[90] D. Blackie, R. Blackwell-Whitehead, G. Stark, J. C. Pickering, P. L. Smith, J. Rufus, and A. P. Thorne. High-resolution photoabsorption cross-section measurements of SO2 at 198 K from 213 to 325 nm. *Journal of Geophysical Research*, 116:E03006, 2011. Erratum, *ibid* (2011) 116, e12099.

[91] R. Kullmer and W. Demtroder. Vibronic coupling in SO2, and its influence on the rotational structure of the bands in the 300-330 nm region. *Chemical Physics*, 92:423–433, 1985.

[92] S. O. Danielache, S. Hattori, M. S. Johnson, Y. Ueno, S. Nanbu, and N. Yoshida. Photoabsorption cross-section measurements of S-32, S-33, S-34, and S-36 sulfur dioxide for the B^1B_1 - X^1A_1 absorption band. *Journal of Geophysical Research*, 117:D24301, 2012.

[93] J.C.D. Brand and R. Nanes. The 3000-3400 Å absorption of sulfur dioxide. *Journal of Molecular Spectroscopy*, 46:194–199, 1973.

[94] Y. Hamada and A. J. Merer. Rotational Structure in the Absorption Spectrum of SO2 between 3000Å and 3300 Å. *Canadian Journal of Physics*, 53:2555–2576, 1975.

[95] C. Hermans, A. C. Vandaele, and S. Fally. Fourier transform measurements of SO2 absorption cross sections: I. Temperature dependence in the 24 000-29 000 cm(-1) (345-420 nm) region. *J. Quant. Spectrosc. Radiat. Transf.*, 110:756–765, 2009.

[96] H. Gerding and N.F. Moerman. Raman spectrum of the ice-like form of sulphur trioxide. *Nature*, 137:1033–1034, 1936.

[97] H. Gerding and J. Lecomte. Infra-Red Absorption Spectrum of Sulphur Trioxide. *Nature*, 142:718, 1938.

[98] K. J. Palmer. The electron diffraction investigation of sulfur monochloride, sulfur dichloride, sulfur trioxide, thionyl chloride, sulfuryl chloride, vanadium oxytrichloride, and chromyl chloride. *Journal of the American Chemical Society*, 60:2360–2369, 1938.

[99] J. Ortigoso, R. Escribano, and A. G. Maki. The ν_2 and ν_4 IR bands of SO_3. *J. Mol. Spectrosc.*, 138:602–613, 1989.

[100] J. Barber, E. T. H. Chrysostom, T. Masiello, J. W. Nibler, A. Maki, A. Weber, T. A. Blake, and R. L. Sams. Analysis of ν_2, ν_4 infrared hot bands of $^{32}S^{16}O_3$: Resolution of the puzzle of the ν_1 CARS spectrum. *J. Mol. Spectrosc.*, 216:105–112, 2002.

[101] S. W. Sharpe, T. A. Blake, R. L. Sams, A. Maki, T. Masiello, J. Barber, N. Vulpanovici, J. W. Nibler, and A. Weber. The ν_3 and $2\nu_3$ bands of $^{32}S^{16}O_3$, $^{32}S^{18}O_3$, $^{34}S^{16}O_3$ and $^{34}S^{18}O_3$. *J. Mol. Spectrosc.*, 222:142–152, 2003.

[102] A. Kaldor, A. G. Maki, A. J. Dorney, and I. M. Mills. Assignment of ν_2 and ν_4 of SO_3. *J. Mol. Spectrosc.*, 45:247–252, 1973.

[103] E. T. H. Chrysostom, N. Vulpanovici, T. Masiello, J. Barber, J. W. Nibler, A. Weber, A. Maki, and T. A. Blake. Coherent Raman and infrared studies of sulphur trioxide. *J. Mol. Spectrosc.*, 210:233–239, 2001.

[104] A. Maki, T. A. Blake, L. S. Sams, N. Vulpanovici, J. Barber, E. T. H. Chrysostom, T. Masiello, J. W. Nibler, and A. Weber. High-resolution infrared spectra of the ν_2, ν_3, ν_4, and $2\nu_3$ bands of $^{32}S^{16}O_3$. *J. Mol. Spectrosc.*, 210:240–249, 2001.

[105] A. Maki, T. A. Blake, R. L. Sams, J. Frieh, J. Barber, T. Masiello, E. T. H. Chrysostom, J. W. Nibler, and A. Weber. Analysis of some combination-overtone infrared bands of $^{32}S^{16}O_3$. *J. Mol. Spectrosc.*, 225:109–122, 2004.

[106] V. Meyer, D. H. Sutter, and H. Dreizler. The centrifugally induced pure rotational spectrum and the structure of sulfur-trioxide - A microwave fourier-transform study of a non polar molecule. *Z. Naturforsch. A*, 46:710–714, 1991.

[107] E. Kauppi and L. Halonen. A simple curvilinear internal coordinate model for vibrational energy levels. *The Journal of Chemical Physics*, 96:2933, 1992.

[108] S. Cartera, I. M. Millsa, J. N. Murrellb, and A. J. C. Varandasc. Analytical potentials for triatomic molecules. *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, 45:1053–1066, 1982.

[109] A. B. McCoy, D. C. Burleigh, and E. L. Sibert III. Rotationvibration interactions in highly excited states of SO2 and H2CO. *J. Chem. Phys.*, 95:7449, 1991.

[110] R. C. Mayrhofer and E. L. Sibert III. Investigating optimal coordinates for describing vibrational motion. *Theoretica Chimica Acta*, 92:107–122, 1995.

[111] F. Iachello and S. Oss. Overtone frequencies and intensities of bent XY2 molecules in the vibron model. *Journal of Molecular Spectroscopy*, 142:85–107, 1990.

[112] T. Sako and K. Yamanouchi. Algebraic approach to vibrationally highly excited states of SO2. Vibrational wavefunctions from spectroscopy. *Chemical Physics Letters*, 264:403–410, 1997.

[113] Y. Pak and R. C. Woods. Coupled cluster calculations of the potential energy surfaces and spectroscopic constants of SiF2, PF+ 2, SO2, PO 2, and ClO+ 2. *J. Chem. Phys.*, 104:5547, 1996.

[114] J. M. L. Martin. Basis set convergence study of the atomization energy, geometry, and anharmonic force field of SO2: The importance of inner polarization functions. *Journal of Chemical Physics*, 108:2791–2800, 1998.

[115] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. A fifth-order perturbation comparison of electron correlation theories . *Chemical Physics Letters*, 157:479–483, 1989.

[116] O. K. Voitsekhovskaya and O. V. Egorov. The absorption of sulfur dioxide in the terahertz range at temperatures of 300-1200 K. *Moscow University Physics Bulletin*, 68:132–138, 2013.

[117] R. C. Fortenberry, X. Huang, D. W. Schwenke, and T. J. Lee. Limited rotational and rovibrational line lists computed with highly accurate quartic force fields

and ab initio dipole surfaces. *Frontiers in Molecular Vibrational Calculations and Computational Spectroscopy*, 119:76–83, 2014.

[118] X. Huang, D. W. Schwenke, and T. J. Lee. Empirical InfraRed Line Lists for Five SO2 Isotopologues: $^{32/33/34/36}S^{16}O_2$ and $^{32}S^{18}O_2$. *J. Mol. Spectrosc.*, 311:19–24, 2015.

[119] P. Nachtigall, J. Hrusak, O. Bludsky, and S. Iwata. Investigation of the potential energy surfaces for the ground $\tilde{X}^1A_1$ and excited $\tilde{C}^1B_2$ electronic states of SO2. *Chem. Phys. Lett.*, 303:441–446, 1999.

[120] O. Bludsky, P. Nachtigall, J. Hrusak, and P. Jensen. The calculation of the vibrational states of SO2 in the $\tilde{C}^1B_2$ electronic state up to the SO((3)Sigma(-))+O(P-3) dissociation limit. *Chem. Phys. Lett.*, 318:607–613, 2000.

[121] A. J. Dorney, A. R. Hoy, and I. M. Mills. Anharmonic force field for so3. *J. Mol. Spectrosc.*, 45:253–260, 1973.

[122] J. P. Flament, N. Rougeau, and M. Tadjeddine. Vibrational frequencies of simple sulfur oxides: ab initio SQM and MCSCF calculations; comparison. *Chemical Physics*, 167:53–60, 1992.

[123] J. M. L. Martin. A fully ab initio quartic force field of spectroscopic quality for SO3. *Spectrochimica Acta A*, 55:709–718, 1999.

[124] N. Jacquinet-Husson, L. Crepeau, R. Armante, C. Boutammine, A. Chédin, N. A. Scott, C. Crevoisier, V. Capelle, C. Boone, N. Poulet-Crovisier, A. Barbe, A. Campargue, D. Chris Benner, Y. Benilan, B. Bézard, V. Boudon, L. R. Brown, L. H. Coudert, A. Coustenis, V. Dana, V. M. Devi, S. Fally, A. Fayt, J.-M. Flaud, A. Goldman, M. Herman, G. J. Harris, D. Jacquemart, A. Jolly, I. Kleiner, A. Kleinböhl, F. Kwabia-Tchana, N. Lavrentieva, N. Lacome, Li-Hong Xu, O. M. Lyulin, J.-Y. Mandin, A. Maki, S. Mikhailenko, C. E. Miller, T. Mishina, N. Moazzen-Ahmadi, H. S. P. Müller, A. Nikitin, J. Orphal, V. Perevalov, A. Perrin, D. T. Petkie, A. Predoi-Cross, C. P. Rinsland, J. J. Remedios, M. Rotger, M. A. H. Smith, K. Sung, S. Tashkun, J. Tennyson, R. A. Toth, A.-C. Vandaele,

and J. Vander Auwera. The 2009 edition of the GEISA spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.*, 112:2395–2445, 2011.

[125] N. Jacquinet-Husson et al. The 2015 edition of the GEISA spectroscopic database. *J. Quant. Spec. Rad. Trans.*, 112:2395–2445, 2011.

[126] S. W. Sharpe, T. J. Johnson, R. L. Sams, P. M. Chu, G. C. Rhoderick, and P. A. Johnson. Gas-phase databases for quantitative infrared spectroscopy. *Appl. Spectrosc.*, 58:1452–1461, 2004.

[127] H. M. Pickett, R. L. Poynter, E. A. Cohen, M. L. Delitsky, J. C. Pearson, and H. S. P. Müller. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transf.*, 60:883–890, 1998.

[128] L. S. Rothman, I. E. Gordon, R. J. Barber, H. Dothe, R. R. Gamache, A. Goldman, V. I. Perevalov, S. A. Tashkun, and J. Tennyson. HITEMP, the High-Temperature Molecular Spectroscopic Database. *J. Quant. Spectrosc. Radiat. Transf.*, 111:2139–2150, 2010.

[129] J. Tennyson and S. N. Yurchenko. ExoMol: molecular line lists for exoplanet and other atmospheres. *Mon. Not. R. Astr. Soc.*, 425:21–33, 2012.

[130] C. Sousa-Silva, S. N. Yurchenko, and J. Tennyson. A computed room temperature line list for phosphine. *J. Mol. Spec.*, 288:28–37, 2013.

[131] S. N. Yurchenko, W. Thiel, and P. Jensen. Theoretical ROVibrational energies (TROVE): A robust numerical approach to the calculation of rovibrational energies for polyatomic molecules. *J. Mol. Spec*, 245:126–140, 2007.

[132] R. J. Barber, J. Tennyson, G. J. Harris, and R. N. Tolchenov. A high accuracy computed water line list. *Mon. Not. R. Astr. Soc.*, 368:1087–1094, 2006.

[133] S. N. Yurchenko, R. J. Barber, and J. Tennyson. A variationally computed line list for hot NH_3. *Monthly Notices of the Royal Astronomical Society*, 413:1828–1834, 2011.

[134] S. N. Yurchenko, J. Tennyson, J. Bailey, M. D. J. Hollis, and G. Tinetti. Spectrum of hot methane in astronomical objects using a comprehensive computed line list. *Proc. Natl. Acad. Sci. U. S. A.*, 111(26):9379–9383, 2014.

[135] C. Sousa-Silva, A. F. Al-Refaie, J. Tennyson, and S. N. Yurchenko. ExoMol line lists VIII: A Hot Line List for Phosphine. *Mon. Not. R. Astr. Soc.*, 446:2337–2347, 2015.

[136] G. Tinetti, A. Vidal-Madjar, M-C. Liang, J-P. Beaulieu, Y. Yung, S. Carey, R. J Barber, J. Tennyson, I. Ribas, N. Allard, G. E. Ballester, D. K. Sing, and F. Selsis. Water vapour in the atmosphere of a transiting extrasolar planet. *Nature*, 448:169–171, 2007.

[137] J. Tennyson, S. N. Yurchenko, A. F. Al-Refaie, E. J. Barton, K. L. Chubb, P. A. Coles, S. Diamantopoulou, M. N. Gorman, C. Hill, A. Z. Lam, L. Lodi, L. K. McKemmish, Y. Na, A. Owens, O. L. Polyansky, C. Sousa-Silva, D. S. Underwood, A. Yachmenev, and E. Zak. The ExoMol database: molecular line lists for exoplanet and other hot atmospheres. *J. Mol. Spec.*, 2016.

[138] J. Tennyson, P. Barletta, R. E. A. Kelly, L. Lodi, J. J. Munro, and B. C. Silva. DVR3D: a program suite for the calculation of rotation-vibration spectra of triatomic molecules. *Comput. Phys. Commun.*, 163:85–116, 2004.

[139] L. Lodi and J. Tennyson. Theoretical methods for small-molecule ro-vibrational spectroscopy. *J. Phys.B: At. Mol. Opt. Phys.*, 43:133001, 2010.

[140] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. *Ann. Phys.*, 389:457–484, 1927.

[141] D. R. Hartree. The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods. *Math. Proc. Cambridge Philos. Soc.*, 24:89–132, 1928.

[142] J. C. Slater. The theory of complex spectra. *Phys. Rev.*, 34:1293–1322, 1929.

[143] V. Fock. Approximation method for the solution of the quantum mechanical multibody problems. *Z. Physik*, 61:126–148, 1930.

[144] F. Jensen. *Introduction to Computational Chemistry*. WileyBlackwell, London, 2006.

[145] J. Cizek. On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *J. Chem. Phys.*, 45:4256–4266, 1966.

[146] J. Cizek. On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules. *Adv. Chem. Phys.*, 14:35–89, 1969.

[147] M. Aallay and J. Gauss. Approximate treatment of higher excitations in coupled-cluster theory. *J. Chem. Phys.*, 123:214105, 2005.

[148] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.*, 51:2657–2664, 1969.

[149] R. Ditchfield, W. J. Hehre, and J. A. Pople. SelfConsistent MolecularOrbital Methods. IX. An Extended GaussianType Basis for MolecularOrbital Studies of Organic Molecules. *J. Chem. Phys.*, 54:724–728, 1971.

[150] T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90(2):1007–1023, 1989.

[151] L. Vereecken and J. S. Franciscob. Theoretical studies of atmospheric reaction mechanisms in the troposphere. *Chem. Soc. Rev.*, 41:6259–6293, 2012.

[152] C. Eckart. Some studies concerning rotating axes and polyatomic molecules. *Phys. Rev.*, 47:552–558, 1935.

[153] J. Tennyson and B. T. Sutcliffe. The ab initio calculation of the vibration-rotation spectrum of triatomic systems in the close-coupling approach with KCN and H_2Ne as examples. *J. Chem. Phys.*, 77:4061–4072, 1982.

[154] P. R. Bunker and P. Jensen. *Fundamentals of Molecular Symmetry*. IOP Publishing, Bristol, 2004.

[155] D. O. Harris, G. G. Engerholm, and W. Gwinn. Calculation of matrix elements for onedimensional quantummechanical problems and the application to anharmonic oscillators. *JCM*, 43:515, 1965.

[156] Z. Bačić and J. C. Light. Theoretical methods for rovibrational states of floppy molecules. *Annu. Rev. Phys. Chem.*, 40:469–498, 1989.

[157] J. C. Light and T Carrington Jr. Discrete-variable representations and their utilization. *Adv. Phys. Chem.*, 114:263–310, 2000.

[158] H. Wei and T. Carrington. Explicit expressions for triatomic Eckart frames in Jacobi, Radau, and bond coordinates. *J. Chem. Phys.*, 107:2813–2818, 1997.

[159] Z. Sun, N. Lou, and G. Nyman. Time-Dependent Wave Packet Split Operator Calculations on a Three-Dimensional Fourier Grid in Radau Coordinates Applied to the OClO Photoelectron Spectrum. *J. Phys. Chem. A*, 108:9226–9232, 2004.

[160] B. T. Sutcliffe and J. Tennyson. A general treatment of vibration-rotation coordinates for triatomic molecules. *Int. J. Quantum Chem.*, 39:183–196, 1991.

[161] J. Tennyson and B. T. Sutcliffe. Discretisation to avoid singularities in vibration-rotation hamiltonians: a bisector embedding for AB_2 triatomics. *Int. J. Quantum Chem.*, 42:941–952, 1992.

[162] P. M. Morse. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Phys. Rev.*, 34:57, 1929.

[163] I. S. Gradshteyn and I. H. Ryzhik. *Tables of Integrals, Series and Products*. 1980.

[164] S. Viti. PhD Thesis, Applications of High Resolution Coherent Anti-Stokes Raman Scattering Spectroscopy. *University College London*, 1997.

[165] S. N. Yurchenko, R. J. Barber, A. Yachmenev, W. Thiel, P. Jensen, and J. Tennyson. A Variationally Computed T=300 K Line List for NH_3. *J. Phys. Chem. A*, 113:11845–11855, 2009.

[166] E. T. H. Chrysostom. Applications of high resolution coherent anti-stokes raman scattering spectroscopy. *PhD Thesis, Oregon State University*, 2001.

[167] P. Jensen. Local modes in vibrationrotation spectroscopy. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2:494–512, 2012.

[168] J. W. Cooley. An improved eigenvalue corrector formula for solving the Schrödinger equation for central fields. *Math. Comp.*, 15:363–374, 1961.

[169] C. E. Dykstra and D. J. Malik. Derivative numerov-cooley theory. a method for finding vibrational state properties of diatomic molecules. *J. Chem. Phys.*, 87:2806, 1987.

[170] D. S. Underwood, J. Tennyson, and S. N. Yurchenko. An ab initio variationally computed room-temperature line list for SO_3. *Phys. Chem. Chem. Phys.*, 15:10118–10125, 2013.

[171] A. F. Al-Refaie, S. N. Yurchenko, A. Yachmenev, and J. Tennyson. ExoMol line lists VIII: A variationally computed line-list for hot formaldehyde. *Mon. Not. R. Astr. Soc.*, 448:1704–1714, 2014.

[172] S. N. Yurchenko and J. Tennyson. ExoMol line lists IV: The rotation-vibration spectrum of methane up to 1500 K. *Mon. Not. R. Astr. Soc.*, 440:1649–1661, 2014.

[173] P Jensen. An introduction to the theory of local mode vibrations. *Mol. Phys.*, 98:1253–1285, 2000.

[174] G. Herzberg. Infrared and raman spectra of polyatomic molecules. *D. Van Nostrand Company Inc.*, 1995.

[175] J. Tennyson, C. Hill, and S. N. Yurchenko. Data structures for ExoMol: Molecular line lists for exoplanet and other atmospheres. In J. D. Gillaspy, W. L. Wiese, and Y. A. Podpaly, editors, *Eighth International Conference on Atomic and Molecular Data and their Applications (ICAMDATA-2012)*, volume 1545 of *AIP Conference Proceedings*, pages 186–195. 2013.

[176] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes – The Art of Scientific Computing*. Cambridge University Press, 3rd edition, 2007.

[177] A. H. Stroud and D. Secrest. *Gaussian Quadrature Formulas*. Prentice-Hall, London, 1966.

[178] J. Ramanlal. Phd thesis, the spectroscopy of h+3: Low energy to dissociation. *University College London*, 2005.

[179] H. Partridge and D. W. Schwenke. The determination of an accurate isotope dependent potential energy surface for water from extensive ab initio calculations and experimental data. *J. Chem. Phys.*, 106:4618–4639, 1997.

[180] X. Huang, D. W. Schwenke, and T. J. Lee. An accurate global potential energy surface, dipole moment surface, and rovibrational frequencies for NH_3. *J. Chem. Phys.*, 129:214304, 2008.

[181] X. Huang, D. W. Schwenke, and T. J. Lee. Rovibrational spectra of ammonia. i. unprecedented accuracy of a potential energy surface used with nonadiabatic corrections. *J. Chem. Phys.*, 134:044320, 2011.

[182] X. Huang, D. W. Schwenke, S. A. Tashkun, and T. J. Lee. An isotopic-independent highly accurate potential energy surface for CO_2 isotopologues and an initial $^{12}C^{16}O_2$ infrared line list. *J. Chem. Phys.*, 136:124311, 2012.

[183] D. W. Schwenke. Variational calculations of rovibrational energy levels and transition intensities for tetratomic molecules. *J. Phys. Chem.*, 100:2867–2884, 1996.

[184] D. E. Woon and T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.*, 98:1358–1371, 1993.

[185] T. H. Dunning. Gaussian-basis sets for use in correlated molecular calculations. 1. the atoms Boron through Neon and Hydrogen. *J. Chem. Phys.*, 90:1007–1023, 1989.

[186] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

[187] A. Fateev. Personal communication. 2015.

[188] C. Hill, S. N. Yurchenko, and J. Tennyson. Temperature-dependent molecular absorption cross sections for exoplanets and other atmospheres. *Icarus*, 226(2):1673–1677, 2013.

[189] S. N. Yurchenko, M. Carvajal, P. Jensen, H. Lin, J. J. Zheng, and W. Thiel. Rotation-vibration motion of pyramidal XY_3 molecules described in the Eckart frame: Theory and application to NH_3. *J. Mol. Phys.*, 103:359–378, 2005.

[190] T. B. Adler, G. Knizia, and H.-J. Werner. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.*, 127:221106, 2007.

[191] K. E. Yousaf and K. A. Peterson. Auxiliary RI Basis Sets (OptRI) matched to aug-cc-pV(n+d)Z: H-He, B-Ne, Al-Ar. *Chem. Phys. Letters*, 476:303, 2009.

[192] H. J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, and M. Schütz. MOLPRO, a package of ab initio programs, 2010. see `http://www.molpro.net/`.

[193] H. Lin, W. Thiel, S. N. Yurchenko, M. Carvajal, and P. Jensen. Vibrational energies for NH_3 based on high level ab initio potential energy surfaces. *J. Chem. Phys.*, 117:11265–11276, 2002.

[194] S. N. Yurchenko, W. Thiel, and P. Jensen. Rotational energy cluster formation in XY_3 molecules: Excited vibrational states of BiH_3 and SbH_3. *J. Mol. Spec.*, 240:174–187, 2006.

[195] R. I. Ovsyannikov, W. Thiel, S. N. Yurchenko, M. Carvajal, and P. Jensen. PH_3 revisited: Theoretical transition moments for the vibrational transitions below 7000 cm^{-1}. *J. Mol. Spec.*, 252:121–128, 2008.

[196] S. N. Yurchenko, W. Thiel, M. Carvajal, and P. Jensen. Ab initio potential energy surface, electric-dipole moment, polarizability tensor, and theoretical rovibrational spectra in the electronic ground state of $^{14}NH_3^+$. *J. Chem. Phys.*, 346:146–159, 2008.

[197] A. V. Nikitin, M. Rey, and V. G. Tyuterev. Rotational and vibrational energy levels of methane calculated from a new potential energy surface. *Chem. Phys. Letters*, 501(4-6):179–186, 2011.

[198] S. N. Yurchenko, R. J. Barber, and J. Tennyson. A variationally computed hot line list for NH_3. *Mon. Not. R. Astr. Soc.*, 413:1828–1834, 2011.

[199] A. E. Lynas-Gray, S. Miller, and J. Tennyson. Infra red transition intensities for water: a comparison of *ab initio* and fitted dipole moment surfaces. *J. Mol. Spectrosc.*, 169:458–467, 1995.

[200] L. Lodi, J. Tennyson, and O. L. Polyansky. A global, high accuracy ab initio dipole moment surface for the electronic ground state of the water molecule. *J. Chem. Phys.*, 135:034113, 2011.

[201] S. N. Yurchenko, J. Tennyson, R. J. Barber, and W. Thiel. Vibrational transition moments of CH_4 from the first principles. *J. Mol. Spectrosc.*, 291:69–76, 2013.

[202] P. Jensen and P. R. Bunker. *Molecular Symmetry and Spectroscopy*. NRC, Canada, 1998.

[203] A. J. Dorney and J. K. G. Watson. Forbidden rotational spectra of polyatomic-molecules: Stark effects and delta J = 0 transitions of Td molecules. *J. Mol. Spectrosc.*, 42:135, 1972.

[204] W. G. Harter and C. W. Patterson. Bands, clusters, and crystal-field splitting - Understanding SF6 rotational levels. *Int. J. Quantum Chem.*, 12:479–492, 1977.

[205] C. W. Patterson and W. G. Harter. Orbital level splitting in octahedral symmetry and SF6 rotational spectra .2. Quantitative treatment of high J levels. *J. Chem. Phys.*, 66:4886–4892, 1977.

[206] W. G. Harter and C. W. Patterson. Orbital level splitting in octahedral symmetry and SF6 rotational spectra .1. Qualitative features of high J levels. *J. Chem. Phys.*, 66:4872–4885, 1977.

[207] W. G. Harter and C. W. Patterson. Simple-model for asymptotic level clusters in SF6 rotational spectra. *Phys. Rev. Lett.*, 38:224–227, 1977.

[208] W. G. Harter and C. W. Patterson. Rotational energy surfaces and high-J eigenvalue structure of polyatomic-molecules. *J. Chem. Phys.*, 80:4241–4261, 1984.

[209] B. I. Zhilinskiĭ and I. M. Pavlichenkov. Critical phenomena in rotational spectra. *Zhurnal Eksperimentalnoi Teor. Fiz.*, 92:387–403, 1987.

[210] W. G. Harter. Computer graphical and semiclassical approaches to molecular rotations and vibrations. *Comp. Phys. Rep.*, 8:319–394, 1988.

[211] B. I. Zhilinsky and I. M. Pavlichenkov. Critical effect in rotational spectra of water molecule. *Opt. Spektrosk.*, 64:688–690, 1988.

[212] D. A. Sadovskiĭ and B. I. Zhilinskiĭ. Qualitative-analysis of vibration-rotation hamiltonians for spherical top molecules. *Mol. Phys.*, 65:109–128, 1988.

[213] D. A. Sadovskiĭ, B. I. Zhilinskiĭ, J. P. Champion, and G. Pierre. Manifestation of bifurcations and diabolic points in molecular-energy spectra. *J. Chem. Phys.*, 92:1523–1537, 1990.

[214] I. N. Kozin and I. M. Pavlichenkov. Bifurcation in rotational spectra of nonlinear AB(2) molecules. *J. Chem. Phys.*, 104:4105–4113, 1996.

[215] I. N. Kozin, R. M. Roberts, and J. Tennyson. Symmetry and structure of H_3^+. *J. Chem. Phys.*, 111:140–150, 1999.

[216] C. van Hecke, D. A. Sadovskiĭ, B. I. Zhilinskiĭ, and V. Boudon. Rotational-vibrational relative equilibria and the structure of quantum energy spectrum of the tetrahedral molecule P-4. *Eur. Phys. J. D*, 17:13–35, 2001.

[217] B. I. Zhilinskiĭ. Symmetry, invariants, and topology in molecular models. *Phys. Rep.-Rev. Sec. Phys. Lett.*, 341:85–171, 2001.

[218] S.V. Petrov and B.M. Kozlovskii. Rotational energy surface and quasiclassical analysis of the rotational energy level cluster formation in the ground vibrational state of ph3. *J. Mol. Spectrosc.*, 243(2):245–252, 2007.

[219] W. G. Harter and J. C. Mitchell. Molecular Eigensolution Symmetry Analysis and Fine Structure. *Int. J. Mol. Sci.*, 14:714–806, 2013.

[220] W. G. Harter. Theory of hyperfine and superfine levels in symmetric polyatomic-molecules .2. Elementary cases in octahedral hexafluoride molecules. *Phys. Rev. A*, 24:192–263, 1981.

[221] P. Jensen. An introduction to the theory of local mode vibrations. *Mol. Phys.*, 98(17):1253–1285, 2000.

[222] P. Jensen, G Osmann, and I. N. Kozin. The Formation of Four-fold Rovibrational Energy Clusters in H_2S, H_2Se, and H_2Te. In D. Papousek, editor, *Advanced Series in Physical Chemistry: Vibration-Rotational Spectroscopy and Molecular Dynamics*, volume 9, pages 298–351. World Scientific Publishing Company, Singapore, 1997.

[223] S. N. Yurchenko, W. Thiel, Per Jensen, and P. R. Bunker. Rotation-vibration energy level clustering in the $\tilde{X}^2B_1$ ground electronic state of PH_2. *J. Mol. Spec.*, 239:160–173, 2006.

[224] S. N. Yurchenko, W. Thiel, S. Patchkovski, and P. Jensen. Theoretical evidence for the formation of rotational energy level clusters in the vibrational ground state of PH_3. *Phys. Chem. Chem. Phys.*, 7:573–582, 2005.

[225] S. N. Yurchenko, M. Carvajal, W. Thiel, and P. Jensen. Ab initio dipole moment and theoretical rovibrational intensities in the electronic ground state of PH_3. *J. Mol. Spec.*, 239:71–87, 2006.

[226] S. N. Yurchenko, R. I. Ovsyannikov, W. Thiel, and P. Jensen. Rotation-vibration energy cluster formation in XH_2D and XHD_2 molecules (X = Bi, P, and Sb). *J. Mol. Spec.*, 256:119–127, 2009.

[227] B. I. Zhilinskii, S. Brodersen, and M. Madsen. The pattern of clusters in isolated vibrational components of CF4 and the semiclassical model. *J. Mol. Spectrosc.*, 160:192–216, 1993.

[228] S. Brodersen and B. I. Zhilinskií. The rotational structure of the vibrational-states and substates of symmetry-e in CF4. *J. Mol. Spectrosc.*, 172:303–318, 1995.

[229] S. Brodersen and B. I. Zhilinskií. Transfer of clusters between the vibrational components of CF4. *J. Mol. Spectrosc.*, 169:1–17, 1995.

[230] J. P. Aldridge, H. Filip, H. Flicker, R. F. Holland, R. S. McDowell, N. G. Nereson, and K. Fox. Octahedral fine-structure splittings in v3 of SF6. *J. Mol. Spectrosc.*, 58(1):165–168, 1975.

[231] J. Karczmarek, J. Wright, P. Corkum, and M. Ivanov. Optical centrifuge for molecules. *Phys. Rev. Lett.*, 82:3420–3423, 1999.

[232] P. R. Bunker and P. Jensen. *Molecular Symmetry and Spectroscopy*. NRC Research Press, Ottawa, 2 edition, 1998.

[233] F. S. Pan and T. Oka. Calculated forbidden rotational spectra of H-3+. *Astrophys. J.*, 305:518–525, 1986.

[234] S. Miller, J. Tennyson, B. Follmeg, P. Rosmus, and H. J. Werner. Ab initio investigation of the bound rovibrational states of the electronic ground state of N_2^+—He. *J. Chem. Phys.*, 89:2178–2184, 1988.

[235] S. V. Petrov and B. M. Kozlovskii. Rotational energy surface and quasiclassical analysis of the rotational energy level cluster formation in the ground vibrational state of PH3. *J. Mol. Spectrosc.*, 243:245–252, 2007.

[236] P. Jensen and I. N. Kozin. The potential-energy surface for the electronic ground-state of H2Se derived from experiment. *J. Mol. Spectrosc.*, 160:39–57, 1993.

[237] S. N. Yurchenko, W. Thiel, M. Carvajal, H. Lin, and P. Jensen. Rotation-vibration motion of pyramidal XY_3 molecules described in the Eckart frame:

The calculation of intensities with application to NH_3. *Adv. Quant. Chem.*, 48:209–238, 2005.

[238] P. R. Bunker and P. Jensen. Chirality in rotational energy level clusters. *J. Mol. Spec.*, 228:640–644, 2004.

[239] S. N. Yurchenko, R. J. Barber, J. Tennyson, W. Thiel, and P. Jensen. Towards efficient refinement of molecular potential energy surfaces: Ammonia as a case study. *J. Mol. Spec.*, 268:123–129, 2011.

[240] A. Yachmenev, S. N. Yurchenko, P. Jensen, and W. Thiel. A new "spectroscopic" potential energy surface for formaldehyde in its ground electronic state. *J. Chem. Phys.*, 134:11, 2011.

[241] S. N. Yurchenko and J. Tennyson. ExoMol line lists - IV. the rotation-vibration spectrum of methane up to 1500 K. *MNRAS*, 440(2):1649–1661, 2014.

[242] R. P. Feynman. Forces in molecules. *Phys. Rev.*, 56:340–343, 1939.

[243] A. F. Al-Refaie, J. Tennyson, and S. N. Yurchenko. GPU Acelerated INtensities (GAIN): a new method of computing Einstein A coefficients. 2015. (to be submitted).

[244] E. J. Barton, S. N. Yurchenko, J. Tennyson, S. Clausen, and A. Fateev. High-resolution absorption measurements of NH_3 at high temperatures: 500 - 2100 cm^{-1}. *J. Quant. Spectrosc. Radiat. Transf.*, 167:126–134, 2015.

[245] A. A. A. Azzam, L. Lodi, S. N. Yurchenko, and J. Tennyson. The dipole moment surface for hydrogen sulfide H_2S. *J. Quant. Spectrosc. Radiat. Transf.*, 161:41–49, 2015.

[246] A. I. Pavlyuchko, S. N. Yurchenko, and J. Tennyson. A hybrid variation-perturbation calculation of the ro-vibrational spectrum of nitric acid. *J. Chem. Phys.*, 142:094309, 2015.

[247] P. E. S. Wormer and J. Tennyson. Ab initio SCF calculations on the potential energy surface of potassium cyanide (KCN). *J. Chem. Phys.*, 75:1245–1252, 1981.

[248] J. Tennyson and B. T. Sutcliffe. Ab initio vibrational-rotational spectrum of potassium cyanide: KCN. *Mol. Phys.*, 46:97–109, 1982.

[249] J. Tennyson and A. van der Avoird. Ab initio vibrational-rotational spectrum of potassium cyanide: KCN. II. Large amplitude motions and ro-vibrational coupling. *J. Chem. Phys.*, 76:5710–5718, 1982.

[250] G. Brocks, J. Tennyson, and A. van der Avoird. Ab initio dipole surfaces, vibrationally averaged dipole moments and infrared transition intensities for KCN and LiCN . *J. Chem. Phys.*, 80:3223–3233, 1984.

[251] J. R. Henderson, H. A. Lam, and J. Tennyson. Highly excited vibrational states of the KCN molecule. *J. Chem. Soc.*, 88:3287–3293, 1992.

[252] S. Brunken, H. S. P. Muller, K. M. Menten, M. C. McCarthy, and P. Thaddeus. The Rotational Spectrum of TiO2. *The American Astronomical Society*, 676:1367–1371, 2008.

[253] C. M. Sharp and W. F. Huebner. Molecular equilibrium with condensation. *Astrophysical Journal Supplement Series*, 72:417–431, 1990.

[254] A. H. Nielsen and P. J. H. Woltz. The infrared spectrum of chlorine dioxide. *The Journal of Chemical Physics*, 20:1878, 1952.

[255] J. Ortigoso, R. Escribano, J. B. Burkholder, and W. J. Lafferty. Infrared Spectrum of OClO in the 2000 cm^1 Region: The $2\nu_1$ and $\nu_1 + \nu_3$ Bands. *Journal of Molecular Spectroscopy*, 158:347356, 1993.

[256] I. Yamamura, T. de Jong, T. Onaka, J. Cami, and L. B. F. M. Waters. Detection of warm SO2 gas in oxygen-rich AGB stars. *Astro. Astrophys.*, 341:L9–L12, 1999.

[257] S. Seager. The future of spectroscopic life detection on exoplanets. *PNAS*, 111:12634–12640, 2014.

[258] A. S. Burrows. Spectra as windows into exoplanet atmospheres. *Proc Natl Acad Sci USA*, 111:12601–12609, 2014.

[259] S. Seager and D. Deming. Exoplanet atmospheres. *Annu. Rev. Astron. Astrophys.*, 48:631–672, 2010.

[260] J. Leconte, F. Forget, and H. Lammer. On the (anticipated) diversity of terrestrial planet atmospheres. *Experimental Astronomy*, 40:449–467, 2014.

[261] H. A. Knutson, D. Dragomir, L. Kreidberg, E. M.-R. Kempton, P. R. McCullough, J. J. Fortney, J. L. Bean, M. Gillon, D. Homeier, and A. W. Howard. Hubble space telescope near-ir transmission spectroscopy of the super-earth HD 97658B. *The Astrophysical Journal*, 794, 2014.

[262] L. Kaltenegger, W. G. Henning, and D. D. Sasselov. Detecting volcanism on extrasolar planets. *Astronomical Journal*, 140, 2010.

[263] J. Tennyson and S.N. Yurchenko. The Status of Spectroscopic Data for the Exoplanet Characterisation Missions. *Exp. Astron.*, 2016.

[264] C. Hill, S. N. Yurchenko, and J. Tennyson. Temperature-dependent molecular absorption cross sections for exoplanets and other atmospheres. *Icarus*, 226:1673–1677, 2013.

[265] J. S. Wilzewski, I. E. Gordon, R. V. Kochanova, C. Hill, and L. S. Rothman. H2, He, and CO2 line-broadening coefficients, pressure shifts and temperature-dependence exponents for the HITRAN database. Part 1: SO2, NH3, HF, HCl, OCS and C2H2. *J. Quantitative Spectroscopy and Radiative Transfer*, 168:193206, 2015.

[266] B. Sumpf. Experimental Investigation of the Self-Broadening Coefficients in the $\nu_1 + \nu_3$ Band of SO2 and the $2\nu_2$ Band of H2S. *J Mol Spectrosc*, 181:160–167, 1997.

[267] B. Sumpf. Line intensity and self-broadening investigations in the 19 micrometer ν_2 band of SO2. *Spectrochimica Acta Part A*, 55:931 1939, 1999.

[268] B. Sumpf. Line intensity and self-broadening investigations in the ν_1 and ν_3 bands of SO2. *Journal of Molecular Structure*, 599:39–49, 2000.

[269] B. Sumpf, M. Schone, and H.-D. Kronfeldt. Self- and Air-Broadening in the ν_3 Band of SO2. *J Mol Spectrosc*, 179:137–141, 1996.

[270] B. Sumpf, O. Fleischmann, and H.-D. Kronfeldt. Self-, Air-, and Nitrogen-Broadening in the ν_1 Band of SO2. *J Mol Spectrosc*, 176:127–132, 1996.

[271] V. V. Lazarev, Y. N. Ponomarev, B. Sumpf, O. Fleischmann, J. Waschull, and H. D. et al. Kronfeldt. Noble gas pressure-induced broadening and shift of H2O and SO2 absorption lines. *J Mol Spectrosc*, 173:17793, 1995.

[272] B. Sumpf, O. Fleischmann, J. Waschull, Y. Heiner, and H.-D. Kronfeldt. Noble gas pressure broadening in the ν_1 and ν_3 band of SO2 studied with IR tunable diode laser spectroscopy. *Infrared Phys Technol*, 36:43945, 1995.

[273] B. Sumpf, M. Schne, O. Fleischmann, Y. Heiner, and H.-D. Kronfeldt. Quantum number and temperature dependence of foreign gas-broadening coefficients in the ν_1 and ν_3 bands of SO2 collisions with H2, air, He, Ne, Ar, Kr, and Xe. *J Mol Spectrosc*, 183:6171, 1997.

[274] N. Tasinato, A. P. Charmet, P. Stoppa, S. Giorgianni, and G. Buffa. Spectroscopic measurements of SO2 line parameters in the 9.2μm atmospheric region and theoretical determination of self-broadening coefficients. *J. Chem. Phys.*, 132, 2010.

[275] C. D. Ball, J. M. Dutta, T. M. Goyette, P. Helminger, and F. C. D. Lucia. The pressure broadening of SO2 by N2, O2, He, and H2 between 90 and 500 K. *J Quantit Spectrosc Radiat Transf*, 56:109–117, 1996.

[276] G. Cazzoli and C. Puzzarini. N2-, O2-, H2-, and He-broadening of SO2 rotational lines in the mm-/submm-wave and THz frequency regions: The J and Ka dependence. *J Quantit Spectrosc Radiat Transf,*, 113:1051–1057, 2012.

[277] N. Tasinato, A.P. Charmet, P. Stoppa, S. Giorgianni, and G. Buffa. N2-, O2- and He-collision-induced broadening of sulfur dioxide ro-vibrational lines in the 9.2 μm atmospheric window. *Spectrochim Acta Part A Mol Biomol Spectrosc*, 118:373–379, 2014.

[278] C. S. Krishnaji. Molecular interaction and linewidth of the asymmetric molecule SO2. II. SO2-CO2 Collisions. *J Chem Phys*, 38:101921, 1963.